Guia de apresentação do Amazon EMR

Amazon EMR

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

Amazon EMR: Guia de apresentação do Amazon EMR

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

As marcas comerciais e imagens comerciais da Amazon não podem ser usadas no contexto de nenhum produto ou serviço que não seja da Amazon, nem de qualquer maneira que possa gerar confusão entre os clientes ou que deprecie ou desprestigie a Amazon. Todas as outras marcas comerciais que não são propriedade da Amazon pertencem aos respectivos proprietários, os quais podem ou não ser afiliados, estar conectados ou ser patrocinados pela Amazon.

Table of Contents

Sobre as versões do Amazon EMR 1
Suporte padrão 2
Noções básicas das versões do Amazon EMR 2
Versionamento da versãoversionamentoversionamento 3
Política de suporte 4
Considerações 7
Versões 7.x do Amazon EMR 9
Versões de aplicações nas versões 7.x do Amazon EMR 9
emr-7.9.0 10
emr-7.8.0 48
emr-7.7.0 92
emr-7.6.0 143
emr-7.5.0 219
emr-7.4.0 294
emr-7.3.0 370
emr-7.2.0 446
emr-7.1.0 565
emr-7.0.0 626
Versões de lançamento 6.x do Amazon EMR 696
Versões de aplicações nas versões 6.x do Amazon EMR 697
emr-6.15.0 697
emr-6.14.0 763
emr-6.13.0 826
emr-6.12.0 898
emr-6.11.1 975
emr-6.11.0 1049
emr-6.10.1 1104
emr-6.10.0 1177
emr-6.9.1 1234
emr-6.9.0 1299
emr-6.8.1 1363
emr-6.8.0 1424
emr-6.7.0 1486
emr-6.6.0 1559

emr-6.5.0 1635
emr-6.4.0 1662
emr-6.3.1 1696
emr-6.3.0 1722
emr-6.2.1 1755
emr-6.2.0 1782
emr-6.1.1 1815
emr-6.1.0 1836
emr-6.0.1 1862
emr-6.0.0 1880
Versões de lançamento 5.x do Amazon EMR 1902
Versões de aplicações nas versões 5.x do Amazon EMR 1906
emr-5.36.2 1906
emr-5.36.1 1950
emr-5.36.0 2012
emr-5.35.0 2061
emr-5.34.0 2088
emr-5.33.1 2113
emr-5.33.0 2142
emr-5.32.1 2165
emr-5.32.0 2189
emr-5.31.1 2217
emr-5.31.0 2236
emr-5.30.2 2260
emr-5.30.1 2279
emr-5.30.0 2301
emr-5.29.0 2325
emr-5.28.1 2344
emr-5.28.0 2362
emr-5.27.1 2382
emr-5.27.0 2400
emr-5.26.0 2419
emr-5.25.0 2439
emr-5.24.1 2458
emr-5.24.0 2476
emr-5.23.1 2494

emr-5.23.0 2511
emr-5.22.0 2530
emr-5.21.2 2549
emr-5.21.1 2566
emr-5.21.0 2583
emr-5.20.1 2602
emr-5.20.0 2619
emr-5.19.1 2639
emr-5.19.0 2655
emr-5.18.1 2673
emr-5.18.0 2690
emr-5.17.2 2707
emr-5.17.1 2724
emr-5.17.0 2740
emr-5.16.1 2758
emr-5.16.0 2774
emr-5.15.1 2791
emr-5.15.0 2807
emr-5.14.2 2824
emr-5.14.1 2840
emr-5.14.0 2856
emr-5.13.1 2874
emr-5.13.0 2889
emr-5.12.3 2905
emr-5.12.2 2921
emr-5.12.1 2936
emr-5.12.0 2952
emr-5.11.4 2969
emr-5.11.3 2984
emr-5.11.2 3000
emr-5.11.1 3015
emr-5.11.0 3031
emr-5.10.1 3047
emr-5.10.0 3062
emr-5.9.1 3079
emr-5.9.0 3094

emr-5.8.3 3111
emr-5.8.2 3126
emr-5.8.1 3141
emr-5.8.0 3156
emr-5.7.1 3172
emr-5.7.0 3187
emr-5.6.1 3202
emr-5.6.0 3217
emr-5.5.4 3232
emr-5.5.3 3247
emr-5.5.2 3261
emr-5.5.1 3276
emr-5.5.0 3291
emr-5.4.1 3306
emr-5.4.0 3321
emr-5.3.2 3336
emr-5.3.1 3351
emr-5.3.0 3366
emr-5.2.3 3381
emr-5.2.2 3395
emr-5.2.1 3410
emr-5.2.0 3425
emr-5.1.1 3440
emr-5.1.0 3455
emr-5.0.3 3469
emr-5.0.2 3484
emr-5.0.1 3497
emr-5.0.0 3511
Versões de lançamento 4.x do Amazon EMR 3526
Versões de aplicações nas versões 4.x do Amazon EMR 3528
Diferenças nas versões 3528
emr-4.9.6 3575
emr-4.9.5 3589
emr-4.9.4 3603
emr-4.9.3 3617
emr-4.9.2 3631

emr-4.9.1 3645
emr-4.8.5 3660
emr-4.8.4 3674
emr-4.8.3 3688
emr-4.8.2 3702
emr-4.8.1 3717
emr-4.8.0 3731
emr-4.7.4 3745
emr-4.7.3 3758
emr-4.7.2 3771
emr-4.7.1 3785
emr-4.7.0 3798
emr-4.6.1 3813
emr-4.6.0 3825
emr-4.5.0 3838
emr-4.4.0 3849
emr-4.3.0 3861
emr-4.2.0 3872
emr-4.1.0 3882
emr-4.0.0 3891
Versões 2.x e 3.x da AMI 3898
Criar um cluster 3899
Instalar aplicações 3902
Personalizar configurações 3902
Hive 3910
HBase 3920
Pig 3933
Spark 3939
S3 DistCp 3943
O que há de novo? 3947
Amazon EMR 7.9.0 3947
Amazon EMR 6.15.0 3958
Amazon EMR 5.36.2 3995
Compatibilidade do SigV4 4016
Abordagem para mitigar o CVE-2021-44228 4016

Solução de ação bootstrap do Amazon EMR para Log4j CVE-2021-44228 e CVE-2021-45046 ... 4017
Perguntas frequentes ... 4025
Arquivo ... 4027
Versão 6.14.0 ... 4028
Versão 6.13.0 ... 4062
Versão 6.12.0 ... 4105
Versão 6.11.1 ... 4153
Versão 6.11.0 ... 4199
Versão 6.10.0 ... 4227
Versão 6.9.0 ... 4257
Versão 6.8.0 ... 4293
Versão 6.7.0 ... 4329
Versão 6.6.0 ... 4377
Versão 5.35.0 ... 4428
Versão 5.34.0 ... 4432
Versão 6.5.0 ... 4435
Versão 6.4.0 ... 4437
Versão 5.32.0 ... 4445
Versão 6.2.0 ... 4451
Versão 5.31.0 ... 4459
Versão 6.1.0 ... 4464
Versão 6.0.0 ... 4471
Versão 5.30.1 ... 4476
Versão 5.30.0 ... 4481
Versão 5.29.0 ... 4486
Versão 5.28.1 ... 4488
Versão 5.28.0 ... 4489
Versão 5.27.0 ... 4491
Versão 5.26.0 ... 4494
Versão 5.25.0 ... 4497
Versão 5.24.1 ... 4499
Versão 5.24.0 ... 4501
Versão 5.23.0 ... 4503
Versão 5.22.0 ... 4505
Versão 5.21.1 ... 4508

Versão 5.21.0 4509
Versão 5.20.0 4512
Versão 5.19.0 4515
Versão 5.18.0 4517
Versão 5.17.1 4518
Versão 5.17.0 4518
Versão 5.16.0 4520
Versão 5.15.0 4521
Versão 5.14.1 4522
Versão 5.14.0 4522
Versão 5.13.0 4525
Versão 5.12.2 4525
Versão 5.12.1 4526
Versão 5.12.0 4526
Versão 5.11.3 4528
Versão 5.11.2 4528
Versão 5.11.1 4528
Versão 5.11.0 4529
Versão 5.10.0 4530
Versão 5.9.0 4532
Versão 5.8.2 4534
Versão 5.8.1 4534
Versão 5.8.0 4535
Versão 5.7.0 4537
Versão 5.6.0 4538
Versão 5.5.3 4538
Versão 5.5.2 4539
Versão 5.5.1 4539
Versão 5.5.0 4539
Versão 5.4.0 4541
Versão 5.3.1 4542
Versão 5.3.0 4542
Versão 5.2.2 4543
Versão 5.2.1 4543
Versão 5.2.0 4544
Versão 5.1.0 4545

Versão 5.0.3 4546
Versão 5.0.0 4546
Versão 4.9.5 4548
Versão 4.9.4 4548
Versão 4.9.3 4548
Versão 4.9.2 4549
Versão 4.9.1 4549
Versão 4.8.4 4549
Versão 4.8.3 4550
Versão 4.8.2 4551
Versão 4.8.0 4551
Versão 4.7.2 4552
Versão 4.7.1 4553
Versão 4.7.0 4553
Versão 4.6.0 4555
Versão 4.5.0 4557
Versão 4.4.0 4557
Versão 4.3.0 4559
Versão 4.2.0 4561
Configurar aplicações 4563
Configurar aplicações ao criar um cluster 4565
Fornecer uma configuração no console ao criar um cluster 4566
Fornecer uma configuração usando a AWS CLI ao criar um cluster 4566
Fornecer uma configuração usando o SDK do Java ao criar um cluster 4567
Reconfigurar um grupo de instâncias de um cluster em execução 4567
Considerações ao reconfigurar um grupo de instâncias 4568
Reconfigurar um grupo de instâncias no console 4571
Reconfigurar um grupo de instâncias usando a CLI 4572
Reconfigurar um grupo de instâncias usando o SDK do Java 4577
Solução de problemas 4578
Armazenar dados de configuração sigilosos no AWS Secrets Manager 4581
Criar um segredo 4581
Conceder acesso ao Amazon EMR para recuperar o segredo 4581
Usar o segredo em uma classificação de configuração 4582
Atualizar o valor do segredo 4583
Configurar aplicações para usar uma Máquina Virtual Java específica 4584

Considerações 4584
Substituir a JVM 4586
Portas de serviço 4590
Usuários do aplicativo 4591
Verificar dependências usando o repositório de artefatos 4593
Sistema de arquivos do EMR (EMRFS) 4596
Consistent view 4598
Habilitar a visualização consistente 4602
Noções básicas sobre como a visualização consistente do EMRFS rastreia objetos no Amazon S3 4604
Lógica de novas tentativas 4605
Metadados da visualização consistente do EMRFS 4606
Configurar notificações de consistência para CloudWatch o Amazon SQS 4610
Configurar a visualização consistente 4611
Referência de comando da CLI do EMRFS 4616
Autorizar acesso aos dados do EMRFS no Amazon S3 4628
Criar um provedor de credenciais personalizado para dados do EMRFS no Amazon S3 ... 4628
Gerenciar o AWS Security Token Service endpoint do padrão 4630
Especificar a criptografia do Amazon S3 usando propriedades do EMRFS 4631
Usar AWS KMS keys para criptografia do EMRFS 4632
Criptografia do lado do servidor do Amazon S3 4634
Criptografia do lado do cliente do Amazon S3 4636
Criptografia do lado do cliente Amazon S3 com S3A 4645
Configurar CSE-KMS 4646
Configurar CSE-CUSTOM 4646
Propriedades da criptografia do lado do cliente do Amazon S3 com S3A 4648
CloudWatch Agente da Amazon 4650
Criar um cluster 4652
Permissões obrigatórias 4652
Endpoint obrigatório 4653
Criar um cluster 4653
Métricas padrão 4654
Configuração 4657
Amazon EMR 7.1.0 4657
Amazon EMR 7.0.0 4691
Considerações 4696

Histórico 4697
Delta Lake 4702
Introdução 4702
Usar clusters do Delta Lake 4703
Delta Lake com Flink 4703
Delta Lake com Trino 4708
Delta Lake com Spark 4709
Delta Lake com Spark e Glue 4716
Considerações 4717
Histórico 4717
Flink 4720
Criar um cluster com o Flink 4723
Configurar o Flink 4724
Hive e Glue 4724
Arquivo de configuração 4726
Múltiplos nós primários 4727
Tamanho do processo de memória 4728
Tamanho do arquivo de saída de log 4729
Java 11 4730
Trabalhos do Flink 4735
Iniciar uma aplicação YARN do Flink como uma etapa em um cluster de execução prolongada 4735
Enviar o trabalho para uma aplicação Flink existente em um cluster de execução prolongada 4737
Enviar um trabalho transitório do Flink 4738
Shell do Flink Scala 4741
IU do Flink 4741
Escalador automático do Flink 4743
Visão geral 4743
Considerações 4743
Habilitação do escalador automático 4743
Configurações 4745
Otimização dos tempos de reinício 4750
Recuperação local de tarefas 4751
Pontos de verificação incrementais 4752
Recuperação refinada 4753

Reinício combinado 4754
Flink com Zeppelin 4755
Introdução 4755
Pré-requisitos 4755
Configurar o Zeppelin-Flink em um cluster do EMR 4755
Executar trabalhos do Flink com o Zeppelin-Flink em um cluster do EMR 4757
Histórico de versões do Flink 4762
Notas da versão do Flink por versão 4803
Ganglia 4806
Criar um cluster com o Ganglia 4808
Visualizar métricas do Ganglia 4809
Métricas do Hadoop e do Spark no Ganglia 4810
Histórico de versões do Ganglia 4811
Hadoop 4857
Configurar o Hadoop 4859
Configuração da tarefa 4860
Definições de configuração de daemon do Hadoop 5294
Configuração do HDFS 5620
Criptografia transparente no HDFS com o Amazon EMR 5621
Configurar a criptografia transparente do HDFS 5622
Considerações para criptografia transparente do HDFS 5625
Servidor de gerenciamento de chaves do Hadoop 5625
Criptografia transparente do HDFS em clusters do EMR com múltiplos nós principais 5629
Criar ou executar uma aplicação Hadoop 5631
Criar binários usando o Amazon EMR 5631
Processamento de dados com streaming 5634
Processar dados com um JAR personalizado 5639
Ler objetos restaurados 5643
Exemplos 5643
Considerações 5644
Ative o reconhecimento de acesso não uniforme à memória para contêineres do YARN 5645
Embalagem de contêiner YARN 5647
Ative o recurso 5647
Considerações 5648
Histórico de versões do Hadoop 5648
Notas de versão do Hadoop por versão 5714

HBase 5728
Criação de um cluster com HBase 5732
Criação de um cluster HBase usando o console 5732
Criando um cluster HBase usando o AWS CLI 5732
HBase no Amazon S3 (modo de armazenamento Amazon S3) 5733
Habilitando HBase no Amazon S3 5735
Usar um cluster de réplica de leitura 5736
HFile Monitoramento persistente 5738
Acompanhamento de arquivos da loja 5740
Considerações operacionais 5741
Logs de gravação antecipada (WAL) do Amazon EMR 5746
Workspaces do WAL 5747
Permissões obrigatórias 5748
Como habilitar o WAL 5749
Restauração do WAL 5751
Configurações de segurança 5753
Usando AWS PrivateLink 5753
Preços e métricas do WAL 5755
Marcação de Workspaces do WAL 5756
Replicação entre clusters do EMR WAL 5757
Considerações e disponibilidade 5759
Referência da CLI do EMRWAL 5761
Usando a HBase concha 5764
Criar uma tabela 5765
Inserir um valor 5765
Obter um valor 5765
Excluir uma tabela 5765
Acesse HBase tabelas com o Hive 5766
Usando HBase instantâneos 5767
Criar um snapshot usando uma tabela 5768
Excluir um snapshot 5768
Exibir informações do snapshot 5768
Exportar um snapshot para o Amazon S3 5769
Importar um snapshot do Amazon S3 5769
Restaurar uma tabela a partir de instantâneos dentro do shell HBase 5770
Configure HBase 5771

Alterações na alocação de memória do YARN 5772
HBase números de porta 5773
HBase configurações do site para otimizar 5774
Considerações sobre a performance 5776
Visualize a interface HBase do usuário 5777
Exibir arquivos HBase de log 5779
Monitor HBase com Ganglia 5780
Monitorando o EMR com HBase a Amazon CloudWatch 5782
Configurar métricas 5784
Usando o destino de métricas 5786
Migrando de versões anteriores HBase 5787
Migração para o Amazon EMR versão 7.4.0 ou posterior 5787
HBase histórico de lançamentos 5789
HCatalog 5867
Criando um cluster com HCatalog 5868
Usando HCatalog 5869
Desabilitar a gravação direta ao usar HCatalog HStorer 5869
Criar uma tabela usando a HCat CLI e usar esses dados no Pig 5870
Acessar a tabela usando o Spark SQL 5871
Exemplo: criar uma HCatalog tabela e gravar nela usando o Pig 5873
HCatalog histórico de versões 5874
Hive 5946
Diferenças e considerações sobre o Hive no Amazon EMR 5949
Diferenças entre o Apache Hive no Amazon EMR e o Apache Hive 5949
Diferenças no Hive entre as versões 4.x e 5.x do Amazon EMR 5950
Atributos adicionais do Hive no Amazon EMR 5951
Configurar um metastore externo para o Hive 5956
Usar o Catálogo de Dados do AWS Glue como metastore do Hive 5957
Usar um banco de dados externo MySQL ou Amazon Aurora 5964
Usar o driver JDBC do Hive 5967
Melhorar a performance do Hive 5969
Habilitar o confirmador otimizado para EMRFS S3 do Hive 5970
Usando o S3 Select 5971
Otimização do MSCK 5974
Usar o LLAP do Hive 5975
Habilitar o LLAP 5975

Iniciar o LLAP no cluster ... 5976
Verificar o status do LLAP ... 5977
Iniciar ou interromper o LLAP ... 5977
Redimensionar a contagem de daemons do LLAP ... 5977
Criptografia no Hive ... 5978
Criptografia modular em Parquet no Hive ... 5978
Criptografia em trânsito em HS2 ... 5982
Histórico de versões do Hive ... 5983
Notas da versão do Hive por versão ... 6057
Hudi ... 6149
Como o Hudi funciona ... 6151
Noções básicas sobre os tipos de armazenamento de conjuntos de dados: copiar na gravação ou mesclar na leitura ... 6151
Registrar um conjunto de dados do Hudi com seu metastore ... 6153
Considerações e limitações ... 6153
Criar um cluster com o Hudi instalado ... 6155
Trabalhe com um conjunto de dados do Hudi ... 6156
Inicializar uma sessão do Spark para Hudi ... 6160
Gravar em um conjunto de dados do Hudi ... 6160
Upsert dados ... 6165
Excluir um registro ... 6167
Ler em um conjunto de dados do Hudi ... 6168
Usar a CLI do Hudi ... 6170
... 6171
Histórico de versões do Hudi ... 6172
Hue ... 6176
Versões do Hue ... 6176
Atributos compatíveis e incompatíveis do Hue no Amazon EMR ... 6178
Considerações ... 6179
Performance com grandes tabelas de metadados do Hue ... 6179
Incompatibilidade entre versões do Hue ... 6181
Conectar-se à interface de usuário da Web do Hue ... 6181
Usar o Hue com um banco de dados remoto no Amazon RDS ... 6182
Solução de problemas ... 6184
Configurações avançadas do Hue ... 6185
Configurar o Hue para usuários do LDAP ... 6185

Histórico de versões do Hue 6188
Iceberg 6254
Como o Iceberg funciona 6255
Use um cluster Iceberg 6257
Usar um cluster do Iceberg com o Spark 6257
Usar um cluster do Iceberg com o Trino 6266
Usar um cluster do Iceberg com o Flink 6268
Usar um cluster do Iceberg com o Hive 6273
Histórico de versões do Iceberg 6278
Notas da versão do Iceberg por versão 6279
Bloco de anotações Jupyter 6282
EMR Studio 6282
Bloco de anotações do EMR 6282
JupyterHub 6283
Criar um cluster com JupyterHub 6287
Considerações ao usar JupyterHub no Amazon EMR 6289
Configurando JupyterHub 6290
Configurar a persistência de cadernos no Amazon S3 6291
Conectar-se ao nó principal e aos servidores de cadernos 6292
JupyterHub configuração e administração 6293
Adicionar usuários e administradores do Caderno Jupyter 6295
Instalar Kernels e bibliotecas adicionais 6306
JupyterHub histórico de versões 6311
Livy 6351
Habilitar HTTPS 6353
Histórico de versões do Livy 6354
MXNet 6402
MXNet histórico de lançamentos 6404
Oozie 6433
Usar o Oozie com um banco de dados remoto no Amazon RDS 6436
Configurar a versão do Java para Oozie 6438
Histórico de versões do Oozie 6439
Notas de lançamento do Oozie por versão 6500
Phoenix 6503
Criar um cluster com o Phoenix 6506
Personalizar as configurações do Phoenix 6507

Clientes do Phoenix 6508
Histórico de versões do Phoenix 6511
Pig 6593
Enviar um trabalho do Pig 6595
Enviar trabalhos do Pig usando o console do Amazon EMR 6596
Enviar trabalhos do Pig usando a AWS CLI 6597
Chamar funções definidas pelo usuário no Pig 6598
Chamada de arquivos JAR a partir do Pig 6598
Chamar scripts em Python/Jython no Pig 6599
Histórico de versões do Pig 6600
Presto 6669
Usar o Presto com o Catálogo de Dados do AWS Glue 6671
Especificar o Catálogo de Dados do AWS Glue como metastore 6672
permissões do IAM 5961
Considerações ao usar o Catálogo de Dados do AWS Glue 6677
Usando o S3 Select Pushdown 6678
O S3 Select Pushdown é adequado para minha aplicação? 6678
Considerações e limitações 6679
Habilitar o S3 Select Pushdown com o PrestoDB ou Trino 6679
Adicionar conectores de banco de dados 6680
Uso de SSL/TLS e LDAPS 6681
Usar autenticação LDAP 6682
Ativar o modo estrito do Presto 6690
Considerações 6692
Lidar com a perda de instâncias spot no Presto 6692
Usar a escalabilidade automática do Presto com desativação tranquila 6694
Considerações com o Presto no Amazon EMR 6696
Executável da linha de comando do Presto 6696
Propriedades de implantação do Presto não configuráveis 6696
Instalação do PrestoDB 6697
Configuração do EMRFS e do PrestoS3 FileSystem 6697
Configuração padrão para personificação do usuário final 6698
Porta padrão para interface Web do Presto 6698
Problema com a execução do bucket do Hive em algumas versões 6698
Histórico de versões do Presto 6699
Spark 6757

Criação de um cluster do Spark 6761
Execução de aplicações do Spark com o Docker no Amazon EMR 6.x 6764
Considerações 6764
Criar uma imagem do Docker 6765
Usar imagens do Docker do Amazon ECR 6766
Use o catálogo do AWS Glue Data Catalog com o Spark no Amazon EMR 6772
Especificando o AWS Glue Data Catalog como metastore do Apache Hive 6773
Especificando o AWS Glue Data Catalog como o catálogo Apache Iceberg 6775
permissões do IAM 5961
Considerações 5962
Trabalhando com uma hierarquia de vários catálogos no AWS Glue Data Catalog 6778
Como os recursos do catálogo são organizados 6778
Configurando um catálogo de dados para uso com o Amazon EMR 6779
Permissões para acessar recursos no AWS Glue Data Catalog 6779
Configure o Spark para acessar uma hierarquia de vários catálogos no AWS Glue Data Catalog 6780
Considerações e limitações para uma configuração de vários catálogos 6783
Configurar o Spark 6783
Padrões do Spark 6784
Configuração da coleta de resíduos 6786
`maximizeResourceAllocation` 6787
Comportamento de desativação 6789
Variável de ThriftServer ambiente Spark 6793
Alterar as configurações padrão do Spark 6793
Migrar do Apache Log4j 1.x para Log4j 2.x 6796
Otimizar a performance do Spark 6796
Execução adaptável de consultas 6797
Corte de partição dinâmico 6799
Nivelamento de subconsultas escalares 6801
DISTINCT antes de INTERSECT 6802
Junção com filtro Bloom 6803
Reordenação de junção otimizada 6804
Armazenamento em Cache de Fragmentos Resultantes 6804
Habilitar o Armazenamento em Cache de Fragmentos Resultantes do Spark 6805
Considerações 6806
Uso do RAPIDS Accelerator 6808

Selecionar tipos de instância ... 6808
Definição das configurações da aplicação ... 6809
Adição de uma ação de bootstrap ... 6816
Iniciar o cluster ... 6817
Acessar o shell do Spark ... 6817
Use o Amazon SageMaker Spark para aprendizado de máquina ... 6819
Escrever uma aplicação do Spark ... 6820
Scala ... 6820
Java ... 6821
Python ... 6822
Melhora da performance do Spark com o S3 ... 6823
User S3 Select ... 6824
Confirmador otimizado para EMRFS S3 ... 6828
Use o protocolo de confirmação otimizada para EMRFS S3 ... 6835
Tentar novamente solicitações do S3 ... 6842
Adicionar uma etapa do Spark ... 6845
Substituir as definições de configuração padrão do Spark ... 6848
Visualizar o histórico de aplicações do Spark ... 6849
Acesse a web do Spark UIs ... 6849
Uso do Spark com o Amazon Kinesis Data Streams ... 6850
Usar o Spark no Amazon Redshift ... 6850
Iniciar uma aplicação do Spark ... 6851
Autenticação no Amazon Redshift ... 6852
Leitura e gravação para o Amazon Redshift ... 6854
Considerações ... 6856
Histórico de versões do Spark ... 6857
Sqoop ... 6924
Versões do Sqoop ... 6924
Considerações sobre o Sqoop no Amazon EMR ... 6926
Usando o Sqoop com integração HCatalog ... 6926
Suporte do JDBC e de banco de dados para o Sqoop ... 6927
Proteção da senha ... 6928
Histórico de versões do Sqoop ... 6929
TensorFlow ... 6984
TensorFlow compilações por tipo de EC2 instância da Amazon ... 6986
Segurança ... 6986

Usando TensorBoard ... 6986
TensorFlow histórico de versões ... 6987
Tez ... 7012
Criar um cluster com o Tez ... 7014
Configurar o Tez ... 7015
Exemplo de configuração ... 7015
Abertura de divisão assíncrona do Tez ... 7016
Interface do usuário da Web do Tez ... 7019
Timeline Server ... 7020
Histórico de versões do Tez ... 7020
Notas da versão do Tez por versão ... 7066
Trino ... 7077
História e design de Trino ... 7078
Conceitos de arquitetura ... 7078
Conceitos de Trino ... 7080
Conclua as etapas de pré-requisito para usar o Amazon EMR com Trino ... 7080
Inicialização de um cluster do Amazon EMR com o Trino ... 7080
Conectar-se ao nó primário e executar as consultas ... 7083
Configurar o Trino no Amazon EMR ... 7084
Configurando conectores para Trino ... 7084
Monitoramento ... 7085
Melhores práticas para Trino no Amazon EMR ... 7085
Principais áreas de foco para melhoria de desempenho ... 7086
Colete e utilize estatísticas de tabelas ... 7088
Desafios comuns ao escalar as cargas de trabalho do Trino ... 7088
Trino ... 7094
Propriedades de implantação do Trino não configuráveis ... 7094
Considerações adicionais ... 7094
Histórico de versões do Trino ... 7095
Trino por versão ... 7095
Amazon EMR 7.6.0: notas de versão do Trino ... 7095
Amazon EMR 7.3.0: notas de versão do Trino ... 7095
Amazon EMR 6.6.0: notas de versão do Trino ... 7096
Zeppelin ... 7097
Considerações ao usar o Zeppelin no Amazon EMR ... 7099
Histórico de versões do Zeppelin ... 7100

ZooKeeper 7162
ZooKeeper histórico de versões 7164
ZooKeeper notas da versão 7204
Amazon EMR 7.8.0 - Notas da versão ZooKeeper 7205
Amazon EMR 7.5.0: notas da versão ZooKeeper 7205
Amazon EMR 7.4.0: notas da versão ZooKeeper 7205
Conectores e utilitários 7206
Exportar, consultar e unir tabelas no DynamoDB 7206
Configurar uma tabela do Hive para executar comandos do Hive 7208
Exemplos de comandos do Hive para exportar, importar e consultar dados 7217
Otimizar o desempenho 7226
Kinesis 7230
O que fazer com a integração entre Amazon EMR e o Amazon Kinesis? 7230
Análise com ponto de verificação de fluxos do Amazon Kinesis 7231
Considerações sobre a performance 7233
Agendar a análise do Amazon Kinesis com o Amazon EMR 7233
Migração do conector do Spark Kinesis para o SDK 2.x do Amazon EMR 7.0 7233
S3 DistCp (s3-dist-cp) 7242
Opções do S3 DistCp 7243
Adicionando o S3 DistCp como uma etapa em um cluster 7251
Limpando após falhas em trabalhos do S3 DistCp 7253
Executar comandos e scripts em um cluster 7255
Enviar uma etapa JAR personalizada para executar um script ou comando 7256
Outras formas de usar `command-runner.jar` 7257
AWS Glossário 7259
..... 7260

Sobre as versões do Amazon EMR

Uma versão do Amazon EMR é um conjunto de aplicações de código aberto no ecossistema de big data. Cada versão contém diferentes aplicações de big data, componentes e atributos que você seleciona para que o Amazon EMR instale e configure quando você criar um cluster. Os aplicativos são empacotados usando um sistema baseado no Apache BigTop, que é um projeto de código aberto associado ao ecossistema Hadoop. Este guia fornece informações sobre aplicações incluídas nas versões do Amazon EMR.

Para obter mais informações sobre conceitos básicos e como trabalhar com o Amazon EMR, consulte o Guia de gerenciamento do Amazon EMR.

Ao executar um cluster, você pode escolher entre várias versões do Amazon EMR. Isso permite que você teste e use versões de aplicações que atendem aos seus requisitos de compatibilidade. Você especifica o número da versão com o rótulo da versão. Os rótulos de versão estão no formato `emr-x.x.x`. Por exemplo, `.emr-7.9.0`

Desde a versão 5.18.0 do Amazon EMR, você pode usar o repositório de artefatos do Amazon EMR para criar o código de trabalho em comparação com as versões exatas de bibliotecas e dependências disponíveis com versões específicas do Amazon EMR. Para obter mais informações, consulte Verificar dependências usando o repositório de artefatos do Amazon EMR.

Para receber atualizações quando uma nova versão do Amazon EMR estiver disponível, assine o RSS feed das notas de versão do Amazon EMR.

Detalhes da versão mais recente, incluindo versões da aplicação, notas de versão, componentes e classificações de configuração das séries 7.x, 6.x e 5.x do Amazon EMR:

- Amazon EMR versão 7.9.0
- Amazon EMR Release 6.15.0
- Versão 5.36.2 do Amazon EMR

> **Note**
>
> Novas versões do Amazon EMR são disponibilizadas em diferentes regiões durante um período de vários anos, começando com a primeira região na data da versão inicial. A versão mais recente pode não estar disponível em sua região durante esse período.

Notas de versão sobre as versões mais recentes do Amazon EMR e um histórico de todas as versões:

- O que há de novo?
- Arquivo de notas de versão do Amazon EMR

Um histórico abrangente de versões das aplicações em cada versão do Amazon EMR:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Detalhes sobre cada versão do Amazon EMR e diferenças entre séries de versão, quando aplicável:

- Versões 7.x do Amazon EMR
- Versões de lançamento 6.x do Amazon EMR
- Versões de lançamento 5.x do Amazon EMR
- Versões de lançamento 4.x do Amazon EMR
- Versões 2.x e 3.x da AMI do Amazon EMR

Suporte padrão ao Amazon EMR

Noções básicas das versões do Amazon EMR

A versão do Amazon EMR é o veículo que fornece o software necessário para executar seus aplicativos de código aberto em EC2 plataformas EKS ou sem servidor. Uma versão do Amazon EMR é composta por três partes: o Ambiente de runtime, os Mecanismos centrais e partes Extras.

- Ambiente de runtime: o ambiente de runtime inclui o sistema operacional que faz parte da imagem de máquina da Amazon (AMI) ou de outra imagem de contêiner quando você inicia o cluster ou contêiner (por exemplo, Amazon Linux 2023). Também inclui os runtimes da linguagem, como o Amazon Corretto JDK, com outras ferramentas relacionadas que a imagem instala.
- Mecanismos centrais: incluem o software central de código aberto, como o Apache Hive, e os formatos de tabela aberta, como o Apache Iceberg.

- Extras: incluem bibliotecas de conveniência e pacotes Python, como `mariadb-connector-java` e software de código aberto, como o Apache Pig.

A versão de lançamento do Amazon EMR que você seleciona empacota todos esses componentes em uma nova versão das imagens de máquina da Amazon (AMI) ou imagens de contêiner baseadas no Amazon Linux (AL). O Amazon EMR mantém a mesma versão do AL para todos os nós de um cluster, contêiner EKS ou aplicação sem servidor. O Amazon EMR tem como objetivo obter o ambiente de runtime mais recente em até 90 dias após seu lançamento pelo Amazon Linux e novas versões de código aberto dos mecanimsos centrais em até 90 dias a partir do lançamento upstream. Os extras são lançados conforme necessário. Você encontra a lista dos softwares nas notas de lançamento de cada versão do Amazon EMR.

Versionamento da versãoversionamentoversionamento

Para ajudar você a entender o escopo das mudanças em cada versão, o Amazon EMR usa versionamento semântico. O versionamento semântico segue um formato de `<major>.<minor>.<patch>`. Como exemplo, vamos considerar a versão 7.0.0.

O primeiro dígito indica o lançamento da versão principal ou lançamento principal. Os principais lançamentos geralmente trazem mudanças substanciais, melhorias e novos recursos para o ambiente de runtime ou os mecanismos centrais que não são compatíveis com versões anteriores. Por exemplo, o lançamento principal 7.x usa o Amazon Linux 2023 com o Amazon Corretto 17 JDK como padrão. Como resultado, a versão contém várias alterações importantes em comparação com o lançamento principal 6.x, que vem no Amazon Linux 2 e com o Amazon Corretto 8 JDK como padrão.

O segundo dígito indica o lançamento da versão secundária ou lançamento secundário. Os lançamentos secundários não são compatíveis com versões anteriores e contêm alterações incrementais, melhorias e recursos nos mecanismos centrais e componentes extras, além de novas funcionalidades. Por exemplo, a versão 6.15 usa o Apache Spark 3.4.1, em comparação com a 6.11.1, que vem com o Apache Spark 3.3.2.

O terceiro dígito indica o lançamento de uma versão de patch ou lançamento de patch. Os lançamentos de patch devem ser compatíveis com versões anteriores que contêm correções e atualizações de patches para o mecanismo central e componentes extras, mas sem novas funcionalidades ou atualizações de versão do OSS. Por exemplo, as versões 6.11.1 e 6.11.0 contêm a mesma versão de OSS do Apache Spark 3.3.2. A partir da versão 6.6.x, o lançamento mais recente do patch oferece o patch mais recente do ambiente de runtime disponível. O ambiente de

runtime não afeta o lançamento de patches do Amazon EMR. Por exemplo, a versão 6.11.1 foi lançada com a versão 2.0.20240109.0 do AL de 19 de dezembro de 2023 a 10 de janeiro de 2024, mas com a versão 2.0.20240124.0 do AL a partir de 11 de janeiro de 2024. Algumas versões do Amazon EMR no EKS permitem adicionar tags ou sufixos às versões semânticas para fornecer funcionalidades alternativas. Por exemplo, a versão `emr-6.15.0-latest` do Amazon EMR no EKS lança uma versão habilitada para o Amazon Corretto 8 JDK, enquanto o lançamento `emr-6.15.0-java17-latest` do Amazon EMR no EKS permite lançar uma versão habilitada para o Amazon Corretto 17 JDK.

Política de suporte

Introdução

O Amazon EMR tem como objetivo emitir novas versões secundárias no mínimo uma vez a cada 90 dias e fornecer suporte para versões secundárias por um período de 24 meses a partir da data de lançamento. Esse suporte abrange o ambiente de runtime e os mecanismos centrais com suas dependências associadas, e não abrange extras ou suas dependências associadas. O Runtime Environment também inclui componentes proprietários da plataforma e APIs que são necessários para iniciar, parar e operar os clusters e os aplicativos. Esse cronograma de lançamento consistente garante um ciclo previsível, tornando mais conveniente planejar, testar e fazer a transição para uma versão que forneça suporte e segurança.

O que esperar do suporte padrão

O suporte padrão fornece correções em tíquetes de suporte técnico para problemas que você encontra nos componentes do ambiente de runtime e dos mecanismos centrais nas configurações recomendadas. Todas as correções estão sujeitas à disponibilidade. A configuração recomendada é o uso do Amazon EMR sem nenhuma modificação, adições ou alteração em binários e configurações presentes em uma versão do Amazon EMR, exceto aqueles encontrados na documentação do Amazon EMR. O Amazon EMR implanta correções no patch mais recente, na versão secundária ou principal do Amazon EMR dentro de 90 dias após a verificação da correção. O Amazon EMR aplica correções automaticamente quando você inicia um novo EMR no EC2 cluster, lança um novo Amazon EMR no contêiner EKS ou aciona uma nova tarefa do EMR Serverless. Os componentes extras são fornecidos por conveniência e o Amazon EMR não fornece correções relacionadas a extras.

Componentes

O suporte padrão abrange os componentes do ambiente de runtime e dos mecanismos centrais, por exemplo, sistema operacional, runtimes de linguagem e software central de código aberto, como Apache Hive e Apache Iceberg. Você pode encontrar a lista completa dos componentes compatíveis com cada versão no Sobre as versões do Amazon EMR Amazon Linux FAQs.

A seguinte lista descreve o suporte que oferecemos para diferentes tipos de componentes no suporte padrão:

- Componentes do ambiente de runtime: os componentes do ambiente de runtime receberão correções nos tíquetes de suporte técnico. As correções são categorizadas como (a) bugs críticos, (b) problemas críticos de corrupção de dados e (c) problemas críticos de segurança. Quando qualificado, o Amazon EMR transferirá as correções do ambiente de runtime para versões mais antigas. Para compatibilidade operacional com os componentes de código aberto do mecanismo central, certos componentes do ambiente de runtime devem permanecer em versões específicas para evitar a interrupção das aplicações do cliente. Nesses componentes, o Amazon EMR depende do código aberto upstream para a disponibilidade de correções. Quando as correções estiverem disponíveis no código aberto, forneceremos a versão estável mais recente dentro de 90 dias após a verificação pelo Amazon EMR.
- Componentes dos mecanismos centrais: os mecanismos centrais fornecem as versões mais recentes de muitos projetos de código aberto, cada um com centenas de bibliotecas de dependências transitivas. Embora as comunidades de código aberto que gerenciam esses projetos tentem resolver problemas e vulnerabilidades e exposições comuns (CVE) conhecidas com frequência, as versões mais recentes ainda podem conter bugs conhecidos e. CVEs O Amazon EMR depende do código aberto upstream para disponibilidade de correções e fornecerá a versão estável mais recente como parte dos componentes do mecanismo central em até 90 dias após a verificação pelo Amazon EMR. Em certos casos, o Amazon EMR pode fornecer uma correção para uma CVE em um dos componentes do mecanismo central que precisa ser resolvida antes do código aberto upstream. O Amazon EMR também fornece suporte técnico e correções em recursos adicionados ao código aberto nas configurações recomendadas. Não transferimos correções de componentes do mecanismo central para versões de patch ou secundárias mais antigas.
- Componentes extras: o Amazon EMR não oferece suporte a componentes extras. Componentes extras são projetos de código aberto fornecidos por conveniência, e o Amazon EMR não fornece correções para problemas encontrados neles. Quaisquer solicitações de suporte ou correções podem ser resolvidas por meio da comunidade de código aberto que oferece suporte a esses componentes.

Ciclo de vida do suporte padrão

A seguir, descrevemos as etapas do ciclo de vida do suporte padrão:

- Suporte padrão: as versões do Amazon EMR estão qualificadas para o suporte padrão 24 meses a partir da data de lançamento. Você pode criar tíquetes de suporte técnico e esperar atualizações para os problemas encontrados nesses lançamentos.
- Fim do suporte: após o término do suporte padrão, as versões do Amazon EMR entram no estágio de fim do suporte (EoS) por 12 meses. As versões do EoS não estão qualificadas para suporte técnico e você não poderá criar nenhum ticket para clusters, contêineres ou trabalhos em execução nessas versões. As versões do EoS não receberão correções, patches ou atualizações. As versões do EoS serão removidas do console, mas continuarão disponíveis por meio da API e da AWS CLI. Você sempre pode continuar executando workloads nas versões do EoS. Recomendamos fortemente que você migre para a versão mais recente do Amazon EMR para continuar recebendo patches de segurança, permancer qualificado para receber suporte técnico e criar tickets de suporte quando necessário.
- Fim da vida útil: após o prazo de fim do suporte (EoS), as versões são consideradas de fim da vida útil (EoL). Embora você possa continuar executando clusters de EoL, o Amazon EMR se reserva o direito de remover as versões de EoL da API e do SDK com base em questões case-by-case operacionais e de segurança. É altamente recomendável que você migre para a versão mais recente do Amazon EMR, pois as versões de EoL podem ser removidas da API e do SDK em casos excepcionais.

Suporte intermediário

O Amazon EMR anunciou essa nova política de suporte em 25 de julho de 2024. De acordo com essa política, as versões do Amazon EMR lançadas até 24 de julho de 2022 agora são designadas como fim da vida útil. No entanto, para que você tenha mais tempo para planejar e migrar para versões mais novas, o Amazon EMR oferecerá o suporte intermediário, equivalente ao suporte padrão, para essas versões mais antigas lançadas no período de dois anos antes deste anúncio.

Depois de 25 de julho de 2024, você pode consultar o status atual e os cronogramas de suporte nas notas de versão.

A tabela a seguir mostra o status do suporte para todas as versões existentes do Amazon EMR no momento do anúncio da política, em 25 de julho de 2024:

Lançamentos e períodos de suporte

Versão de lançamento do Amazon EMR	Data da versão inicial	Data de término do suporte padrão	Data de início do término do suporte	Data de início do fim da vida útil
7.2.0	25 de julho de 2024	24 de julho de 2026	25 de julho de 2026	25 de julho de 2027
7.1.0	23 de abril de 2024	22 de abril de 2026	23 de abril de 2026	23 de abril de 2027
7.0.0	19 de dezembro de 2023	18 de dezembro de 2025	19 de dezembro de 2025	19 de dezembro de 2026
5.36.x e de 6.6.x a 6.15.x	De 9 de maio de 2022 a 13 de novembro de 2023	Suporte intermediário até 24 de janeiro de 2026	25 de janeiro de 2026	25 de janeiro de 2026
• Série 6.x: 6.5.0 e inferior • Série 5.x: 5.35.0 e inferior • Séries 4.x, 3.x e 2.x	De 1 de janeiro de 2013 a 30 de março de 2022	Suporte intermediário até 24 de julho de 2025	25 de julho de 2025	25 de julho de 2025

Considerações

O Standard Support está disponível para todos os modelos de implantação do Amazon EMR (EMR on, EC2 Amazon EMR no EKS e EMR Serverless), em todas as regiões onde o Amazon EMR está disponível, sem custo adicional. Os clusters executados com as configurações recomendadas são automaticamente qualificados para suporte, conforme descrito na política, portanto, você não precisa realizar nenhuma ação adicional para ativar o suporte.

- O suporte padrão é compatível somente com componentes necessários para clusters do Amazon EMR. O Amazon EMR não pode garantir a disponibilidade de patches e correções de segurança

caso os componentes de código aberto do mecanismo central cheguem ao EoL upstream ou quando as atualizações de segurança não estiverem mais disponíveis para dependências. Embora você possa optar por instalar os componentes extras, o Amazon EMR não oferecerá suporte a eles ou a suas dependências. Por exemplo, você pode instalar aplicações de terceiros na AMI personalizada para reforçar a segurança do cluster, instalar componentes adicionais ou copiar objetos usando scripts de ação de bootstrap ou SSH no cluster e atualizar as versões padrão do pacote. O Amazon EMR não é compatível com esses componentes. O Standard Support não cobre ações de bootstrap, pacotes, bibliotecas, seu código personalizado e aplicativos bring-your-own personalizados fornecidos pelo cliente que você pode configurar o Amazon EMR para instalar para sua conveniência.

- Seus clusters existentes não serão afetados, independentemente da versão do Amazon EMR que estejam executando. Você pode continuar executando os clusters existentes sem interrupções. Você também pode continuar lançando novos clusters e executando trabalhos em qualquer uma das versões existentes e novas. Todos os lançamentos existentes e novos no momento em que a política entrar em vigor são cobertos pelo suporte padrão por 24 meses a partir da data inicial de lançamento da versão do Amazon EMR. O Amazon EMR fornecerá a você um suporte intermediário durante o anúncio inicial da política. Para receber suporte ininterrupto, recomendamos testar suas aplicações e fazer o upgrade imediato para a versão mais recente do Amazon EMR.
- O Amazon EMR não alterará os componentes do suporte padrão em versões ou clusters existentes. No entanto, o Amazon EMR se reserva o direito de honrar o fim da vida útil inicial e remover esses componentes nas novas versões ou transferir a versão existente para o status End of Support (EoS) ou End of Life (EoL) em casos excepcionais. case-by-case Notificaremos você sobre qualquer remoção por meio dos canais disponíveis.
- Quando uma nova região for lançada, o Amazon EMR só oferecerá suporte às versões de lançamento do Amazon EMR sob o suporte padrão, lançadas seis meses antes da data em que a nova região estiver disponível ao público geral.
- O Amazon EMR não atualizará automaticamente os clusters existentes para as versões mais recentes. No entanto, você pode optar por atualizar os novos clusters para as versões mais recentes do patch, se desejar.
- O custo do uso não suportado do Amazon EMR contará para AWS sua fatura. Mesmo que você esteja usando o Amazon EMR sem suporte, os custos associados a esse uso ainda fazem parte do seu AWS consumo geral e serão incluídos no cálculo das taxas de suporte.

Para obter mais informações, entre em contato com o AWS Developer Support.

Versões 7.x do Amazon EMR

Esta seção contém versões de aplicações, notas da versão, versões de componentes e classificações de configuração disponíveis em cada versão 7.x do Amazon EMR.

Ao executar um cluster, você pode escolher entre várias versões do Amazon EMR. Isso permite que você teste e use versões de aplicações que atendem aos seus requisitos de compatibilidade. Você especifica o número da versão com o rótulo da versão. Os rótulos de versão estão no formato `emr-x.x.x`. Por exemplo, `.emr-7.9.0`

Novas versões do Amazon EMR são disponibilizadas em diferentes regiões durante um período de vários anos, começando com a primeira região na data da versão inicial. A versão mais recente pode não estar disponível em sua região durante esse período.

Para obter uma tabela abrangente das versões de aplicações em cada versão 7.x do Amazon EMR, consulte Versões de aplicações nas versões 7.x do Amazon EMR.

Tópicos

- Versões de aplicações nas versões 7.x do Amazon EMR
- Amazon EMR versão 7.9.0
- Amazon EMR versão 7.8.0
- Amazon EMR versão 7.7.0
- Amazon EMR versão 7.6.0
- Amazon EMR versão 7.5.0
- Amazon EMR versão 7.4.0
- Versão 7.3.0 do Amazon EMR
- Versão 7.2.0 do Amazon EMR
- Versão 7.1.0 do Amazon EMR
- Versão 7.0.0 do Amazon EMR

Versões de aplicações nas versões 7.x do Amazon EMR

Para obter uma tabela abrangente que liste as versões das aplicações disponíveis em cada versão 7.x do Amazon EMR, abra Versões de aplicações nas versões 7.x do Amazon EMR no seu navegador.

Amazon EMR versão 7.9.0

Ciclo de vida compatível com a versão 7.9.0

A tabela a seguir descreve as datas de ciclo de vida suportadas para a versão 7.9.0 do Amazon EMR.

Fase de suporte	Data
Data da versão inicial	19 de maio de 2025
Suporte padrão até	19 de maio de 2027
Término do suporte	20 de maio de 2027
Fim da vida útil	19 de maio de 2028

Componentes da versão 7.9.0

O Amazon EMR 7.9.0 inclui os seguintes aplicativos:

Mecanismos centrais:

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

Extras:

Hue, Oozie JupyterHub, Porco, Presto, Zeppelin TensorFlow

Para obter mais detalhes sobre os pacotes incluídos, consulte Versões do aplicativo nas versões 7.x do Amazon EMR. Para obter mais informações sobre os mecanismos centrais e extras, consulte Amazon EMR Standard Support.

Versões do aplicativo 7.9.0

Essa versão inclui os seguintes aplicativos: AmazonCloudWatchAgentDeltaFlink, HBase, HCatalogHadoop, Hive, Hudi, Hue, IcebergJupyterEnterpriseGateway, JupyterHub, Livy, Oozie, PhoenixPig, Presto, Spark, TensorFlow, TezTrino, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-7.9.0	emr-7.8.0	emr-7.7.0	emr-7.6.0
AWS SDK for Java	2.31.16, 1.12.782	2.29.52, 1.12.780	2.29.44, 1.12.780	2.29.25, 1.12.779
Python	3,9, 3,11	3,9, 3,11	3,9, 3,11	3,9, 3,11
Scala	2.12.18	2.12.18	2.12.18	2.12.18
AmazonCloudWatchAgent	1.300032.2	1.300032.2	1.300032.2	1.300032.2
Delta	3.3.0-amzn-1	3.3.0-amzn-0	3.2.1-amzn-2	3.2.1-amzn-1
Flink	1.20.0	1.20.0	1.20.0	1.20.0
HBase	2.6.2	2.6.1	2.6.1	2.6.1
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.4.1	3.4.1	3.4.0	3.4.0
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-aman-6	0.15.0-amzn-5	0.15.0-aman-4	0.15.0-aman-3
Hue	4.11.0	4.11.0	4.11.0	4.11.0

	emr-7.9.0	emr-7.8.0	emr-7.7.0	emr-7.6.0
Iceberg	1.7.1-amzn-2	1.7.1-amzn-1	1.7.1-amzn-0	1.6.1-amzn-2
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.8.0
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.2.1	5.2.1	5.2.1	5.2.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,287	0,287	0,287	0,287
Spark	3.5.5	3.5.4	3.5.3	3.5.3
TensorFlow	2.16.1	2.16.1	2.16.1	2.16.1
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	467	467	467	457
Zeppelin	0.11.1	0.11.1	0.11.1	0.11.1
ZooKeeper	3.9.3	3.9.3	3.9.2	3.9.2

Notas de lançamento da versão 7.9.0

As notas de lançamento a seguir incluem informações sobre a versão 7.9.0 do Amazon EMR.

Novos recursos

- Atualizações da aplicação: Amazon EMR 7.9.0 application upgrades include AWS SDK for Java 2.31.16/1.12.782, Delta 3.3.0-amzn-1, Hadoop 3.4.1, Hudi 0.15.0-amzn-6, Iceberg 1.7.1-amzn-2, Spark 3.5.5, and Zookeeper 3.9.3.

- Suporte CM-CMK para interface de usuário de aplicativo persistente — Esse recurso estende a criptografia SSE aplicada aos registros gerenciados no bucket de serviços do EMR para honrar a chave mestra personalizada gerenciada pelo cliente (CM-CMK). Esse recurso permite que os clientes gerenciem a acessibilidade dos registros após o encerramento dos clusters.

Problemas conhecidos

- Quando a opção de registro do EMR on EC2 é desativada, os registros não estão disponíveis após o encerramento do cluster e o aplicativo persistente não UIs conterá nenhum registro.

Alterações, melhorias e problemas resolvidos

- Depreciação do Oozie — Como o OSS Oozie entrou no modo sótão, a última versão a incluir o Apache Oozie é a versão 7.9 do Amazon EMR. O Apache Oozie será removido da versão 7.10 e versões superiores do Amazon EMR. Recomendamos que você use o MWAA como alternativa ao Oozie. Os clientes podem consultar o utilitário para GitHub migrar os fluxos oozie-to-airflow-emrde trabalho existentes do Oozie para o MWAA.
- Embalagem de contêineres do YARN — A partir da versão 7.9.0 do Amazon EMR, a política de empacotamento de contêineres agora está disponível para o programador de capacidade do YARN, que se baseia na política de posicionamento de vários nós do YARN.
- O hudi-cli.sh — O script hudi-cli.sh foi descontinuado no EMR versão 7.9.0 e superior. A versão 7.9.0 e superior do Amazon EMR usa .jar. hudi-cli-bundle
- Clusters EMR na AMI AL2 023 habilitada para FIPS — A partir do EMR 7.9, os clientes podem iniciar clusters EMR na AMI 023 habilitada para FIPS. AL2 Essa alteração atualiza as assinaturas digitais e os algoritmos de hash do EMR para garantir que todos os pacotes do EMR sigam as políticas criptográficas recomendadas para o FIPS 140-3. Nota: Isso não abrange o status de certificação FIPS do EMR AL2 ou 023 - consulte as perguntas frequentes do FIPS 023 AL2para obter as informações mais recentes sobre o status de validação do FIPS 140-3.
- Latência de etapas — Esta versão resolve um problema em que a latência de etapas aumentava ao usar criptografia em trânsito sem selecionar explicitamente o aplicativo Hadoop para instalação.
- Lançamento de cluster com volumes criptografados com EBS — Esta versão corrige um problema em que uma chamada KMS em volumes criptografados com EBS pode causar falha na inicialização de um cluster EMR.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250527.1	6.1.134-152.225.amzn2023	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein), China (Pequim), Ásia-Pacífico (Malásia), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Leste dos EUA (Norte da Virgínia), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Malásia), Europa (Londres), Ásia-Pacífico (Melbourne), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (AWS GovCloud AWS GovCloud Oregon), México (Central), Oeste dos EUA (Norte da Califórnia), Ásia-Pacífico

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv), Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250331.(	6.1.131-143.221.am zn2023	18 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central), Ásia-Pacífico (Tóquio)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250303.(	6.1.129-138.220.am zn2023	27 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250218.1	6.1.128-136.201.amzn2023	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderabad), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250211.(	6.1.127-135.201.am zn2023	26 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

Versões Java padrão 7.9.0

O Amazon EMR versão 7.9 e superior vem com o Amazon Corretto 17 (construído no OpenJDK) por padrão para aplicativos que suportam o Corretto 17 (JDK 17), com exceção do Apache Livy.

A tabela a seguir mostra as versões padrão do Java para aplicativos no Amazon EMR 7.9.0. Se você quiser alterar a JVM padrão no cluster, siga as instruções em Configurar aplicações para usar uma Máquina Virtual Java específica para cada aplicação executada no cluster. Você só pode usar uma versão de runtime do Java para um cluster. O Amazon EMR não é compatível com a execução de diferentes nós ou aplicações em diferentes versões de runtime no mesmo cluster.

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
Delta	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Tez	17, 11, 8
Trino	21, 17
Zeppelin	17, 11, 8
Pig	17, 11, 8
Zookeeper	17, 11, 8

Versões do componente 7.9.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *CommunityVersion*-amzn-*EmrVersion*. O *EmrVersion* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado myapp-component com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como 2.2-amzn-2.

Componente	Versão	Descrição
adot-java-agent	1.31.0	Um Java Agent que coleta métricas de daemons de aplicações.
delta	3.3.0-amzn-1	O Delta Lake é um formato de tabela aberto para grandes conjuntos de dados analíticos
emr-amazon-cloudwatch-agent	1.300032.2-amzn-0	Um aplicativo que coleta métricas internas em nível de sistema e métricas de aplicativos personalizadas de instâncias da Amazon. EC2
emr-ddb	5.7.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.18.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.22.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.18.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway

Componente	Versão	Descrição
emr-s3-dist-cp	2.39.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	2.18.0	EMR S3Select Connector
emr-wal-cli	1.3.4	Cli usada para lista/exclusão de emrwal.
emrfs	2.70,0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.20.0-aman-3	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.20.0-aman-3	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
hadoop-client	3.4.1-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	3.4.1-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.4.1-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.4.1-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.

Componente	Versão	Descrição
hadoop-hdfs-journalnode	3.4.1-amzn-1	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.4.1-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.4.1-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	3.4.1-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.4.1-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	3.4.1-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	3.4.1-amzn-1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	2.6.2-amzn-0	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.

Componente	Versão	Descrição
hbase-region-server	2.6.2-amzn-0	Serviço para atender a uma ou mais HBase regiões.
hbase-client	2.6.2-amzn-0	HBase cliente de linha de comando.
hbase-rest-server	2.6.2-amzn-0	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.6.2-amzn-0	Serviço que fornece um endpoint Thrift para. HBase
hbase-operator-tools	2.6.2-amzn-0	Ferramenta de reparo para HBase clusters Apache.
hcatalog-client	3.1.3-aman-18	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.3-aman-18	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	3.1.3-aman-18	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.3-aman-18	O cliente da linha de comando do Hive.
hive-hbase	3.1.3-aman-18	Cliente do Hive-hbase.

Componente	Versão	Descrição
hive-metastore-server	3.1.3-aman-18	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.3-aman-18	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.15.0-aman-6	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.15.0-aman-6	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-trino	0.15.0-aman-6	Biblioteca de pacotes para execução do Presto com o Hudi.
hudi-spark	0.15.0-aman-6	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.11.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
iceberg	1.7.1-amzn-2	Apache Iceberg é um formato de tabela aberta para conjuntos de dados analíticos imensos
jupyterhub	1.5.0	Servidor de vários usuários para blocos de anotações Jupyter

Componente	Versão	Descrição
livy-server	0.8.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.
nvidia-cuda	12.3.0	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitações de fluxo de trabalho do Oozie.
opencv	4.7.0	Biblioteca de Visão Computacional de Código Aberto.
phoenix-library	5.2.1	As bibliotecas do phoenix para servidor e cliente
phoenix-connectors	6.0.0	Conectores do Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica

Componente	Versão	Descrição
presto-coordinator	0,287-amzn-3	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0,287-amzn-3	O serviço que executa partes de uma consulta.
presto-client	0,287-amzn-3	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
trino-coordinator	467-aman-2	Serviço que aceita consultas e gerencia a execução de consultas entre trino-workers.
trino-worker	467-aman-2	O serviço que executa partes de uma consulta.
trino-client	467-aman-2	Cliente de linha de comando do Trino instalado em mestres de espera de um cluster de alta disponibilidade (HA) em que o servidor do Trino não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.3.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger

Componente	Versão	Descrição
spark-client	3.5.5-amzn-0	Os clientes da linha de comando do Spark.
spark-history-server	3.5.5-amzn-0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.5.5-amzn-0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.5.5-amzn-0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
spark-rapids	25.02.1-amzn-0	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs
tensorflow	2.16.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.10.2-amzn-16	O aplicativo e as bibliotecas do YARN para Tez.
tez-on-worker	0.10.2-amzn-16	A aplicação e as bibliotecas do YARN nós de processam ento.
zeppelin-server	0.11.1	O notebook baseado na web que permite um data analytics interativo.

Componente	Versão	Descrição
zookeeper-server	3.9.3-amzn-1	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.9.3-amzn-1	ZooKeeper cliente de linha de comando.

Classificações de configuração 7.9.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

classificações do emr-7.9.0

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j2	Altere os valores no arquivo HBase hbase-log4j2.prope rties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Classificações	Descrição	Ações de reconfiguração
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.

Classificações	Descrição	Ações de reconfiguração
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Alterar a página inicial do KMS do Hadoop	Not available.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.

Classificações	Descrição	Ações de reconfiguração
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.	Not available.
iceberg-defaults	Alterar os valores no arquivo iceberg-defaults.conf do Iceberg.	Not available.
delta-defaults	Alterar os valores no arquivo delta-defaults.conf do Delta.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j2	Alterar as configurações de log4j2.properties do Livy.	Restarts Livy Server.

Classificações	Descrição	Ações de reconfiguração
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j2	Altere os valores no arquivo log4j2.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.

Classificações	Descrição	Ações de reconfiguração
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-lakeformation	Alterar os valores no arquivo lakeformation.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
trino-log	Alterar os valores no arquivo log.properties do Trino.	Restarts Trino-Server (for Trino)
trino-config	Alterar os valores no arquivo config.properties do Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Alterar os valores no arquivo password-authenticator.properties do Trino.	Restarts Trino-Server (for Trino)
trino-env	Alterar os valores no arquivo trino-env.sh do Trino.	Restarts Trino-Server (for Trino)
trino-node	Alterar os valores no arquivo node.properties do Trino.	Not available.
trino-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Trino.	Not available.
trino-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Trino.	Not available.
trino-connector-delta	Alterar os valores no arquivo delta.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Alterar os valores no arquivo hive.properties do Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Alterar os valores no arquivo exchange-manager.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Alterar os valores no arquivo iceberg.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Alterar os valores no arquivo hudi.properties do Trino.	Restarts Trino-Server (for Trino)

Classificações	Descrição	Ações de reconfiguração
trino-connector-jmx	Alterar os valores no arquivo jmx.properties do Trino.	Not available.
trino-connector-kafka	Alterar os valores no arquivo kafka.properties do Trino.	Not available.
trino-connector-localfile	Alterar os valores no arquivo localfile.properties do Trino.	Not available.
trino-connector-memory	Alterar os valores no arquivo memory.properties do Trino.	Not available.
trino-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Trino.	Not available.
trino-connector-mysql	Alterar os valores no arquivo mysql.properties do Trino.	Not available.
trino-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Trino.	Not available.
trino-connector-raptor	Alterar os valores no arquivo raptor.properties do Trino.	Not available.
trino-connector-redis	Alterar os valores no arquivo redis.properties do Trino.	Not available.
trino-connector-redshift	Alterar os valores no arquivo redshift.properties do Trino.	Not available.
trino-connector-tpch	Alterar os valores no arquivo tpch.properties do Trino.	Not available.
trino-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Trino.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Alterar os valores no arquivo kms-logback.xml do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j2	Alterar os valores no arquivo log4j2.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie and HiveServer2.

Classificações	Descrição	Ações de reconfiguração
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.
zookeeper-logback	Altere os valores no ZooKeeper arquivo logback.xml do.	Restarts Zookeeper server.
emr-metrics	Altere as configurações de métricas do EMR desse nó.	Restarts the CloudWatchAgent service.

Registro de alterações 7.9.0

Registro de alterações da versão 7.9.0 e das notas de lançamento

Data	Event	Descrição
2025-05-19	Publicação de documentos	Notas de lançamento do Amazon EMR 7.9.0 publicadas pela primeira vez
2025-05-19	Implantação concluída	Amazon EMR 7.9.0 totalmente implantado em todas as regiões suportadas
2025-05-19	Lançamento inicial limitado	Amazon EMR 7.9.0 foi implantado pela primeira vez nas regiões comerciais iniciais

Amazon EMR versão 7.8.0

Ciclo de vida compatível com 7.8.0

A tabela a seguir descreve as datas de ciclo de vida suportadas para a versão 7.8.0 do Amazon EMR.

Fase de suporte	Data
Data da versão inicial	07 de março de 2025
Suporte padrão até	6 de março de 2027
Término do suporte	7 de março de 2027
Fim da vida útil	7 de março de 2028

Componentes da versão 7.8.0

O Amazon EMR 7.8.0 inclui os seguintes aplicativos:

Mecanismos centrais:

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

Extras:

Hue, Oozie JupyterHub, Porco, Presto, Zeppelin TensorFlow

Para obter mais detalhes sobre os pacotes incluídos, consulte Versões do aplicativo nas versões 7.x do Amazon EMR. Para obter mais informações sobre os mecanismos centrais e extras, consulte Amazon EMR Standard Support.

Versões do aplicativo 7.8.0

Essa versão inclui os seguintes aplicativos: AmazonCloudWatchAgentDeltaFlink, HBase, HCatalogHadoop, Hive, Hudi, Hue, IcebergJupyterEnterpriseGateway, JupyterHub, Livy, Oozie, PhoenixPig, Presto, Spark, TensorFlow, TezTrino, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-7.8.0	emr-7.7.0	emr-7.6.0	emr-7.5.0
AWS SDK for Java	2.29.52, 1.12.780	2.29.44, 1.12.780	2.29.25, 1.12.779	2.28.8, 1.12.772
Python	3,9, 3,11	3,9, 3,11	3,9, 3,11	3,9, 3,11
Scala	2.12.18	2.12.18	2.12.18	2.12.18

	emr-7.8.0	emr-7.7.0	emr-7.6.0	emr-7.5.0
AmazonCloudWatchAgent	1.300032.2	1.300032.2	1.300032.2	1.300032.2
Delta	3.3.0-amzn-0	3.2.1-amzn-2	3.2.1-amzn-1	3.2.1-amzn-0
Flink	1.20.0	1.20.0	1.20.0	1.19.1
Ganglia	-	-	-	-
HBase	2.6.1	2.6.1	2.6.1	2.5.10
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.4.1	3.4.0	3.4.0	3.4.0
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-amzn-5	0.15.0-aman-4	0.15.0-aman-3	0.15.0-amzn-2
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.7.1-amzn-1	1.7.1-amzn-0	1.6.1-amzn-2	1.6.1-amzn-1
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.8.0
MXNet	-	-	-	-
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.2.1	5.2.1	5.2.0	5.2.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0

	emr-7.8.0	emr-7.7.0	emr-7.6.0	emr-7.5.0
Presto	0,287	0,287	0,287	0,287
Spark	3.5.4	3.5.3	3.5.3	3.5.2
Sqoop	-	-	-	-
TensorFlow	2.16.1	2.16.1	2.16.1	2.16.1
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	467	467	457	446
Zeppelin	0.11.1	0.11.1	0.11.1	0.11.1
ZooKeeper	3.9.3	3.9.2	3.9.2	3.9.2

Notas de lançamento da versão 7.8.0

As notas de lançamento a seguir incluem informações sobre a versão 7.8.0 do Amazon EMR.

Novos recursos

- Atualizações da aplicação: Amazon EMR 7.8.0 application upgrades include AWS SDK for Java 2.29.52/1.12.780, Delta 3.3.0-amzn-0, Hadoop 3.4.1, Hudi 0.15.0-amzn-5, Iceberg 1.7.1-amzn-1, Spark 3.5.4, and Zookeeper 3.9.3.

Problemas conhecidos

- Não aplicável a esta versão.

Alterações, melhorias e problemas resolvidos

- Sem grandes mudanças.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250527.1	6.1.134-152.225.amzn2023	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein), China (Pequim), Ásia-Pacífico (Malásia), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Leste dos EUA (Norte da Virgínia), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Malásia), Europa (Londres), Ásia-Pacífico (Melbourne), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (AWS GovCloud AWS GovCloud Oregon), México (Central), Oeste dos EUA (Norte da Califórnia), Ásia-Pacífico

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv), Ásia-Pacífico (Taipei), Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250512.(	6.1.134-152.225.am zn2023	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Hyderabad), Ásia-Pacífico (Jacarta)), Ásia-Pacífico (Malásia), Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Oriente Médio (Oriente Médio) (Oriente Médio (Oriente Médio)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Emirados Árabes Unidos), Canadá (Central) , Oeste do Canadá (Calgary), (Leste dos EUA), AWS GovCloud (AWS GovCloud Oeste dos EUA), China (Pequim), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250428.1	6.1.134-150.224.amzn2023	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Jacarta) Arta), Ásia-Pacífico (Malásia), Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Malásia) , Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (EAU), Canadá (Central) , Oeste do Canadá

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Calgary), (Leste dos EUA), AWS GovCloud (Oeste dos EUA), China (AWS GovCloud Pequim), China (Ningxia), Israel (Tel Aviv), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250414.(	6.1.132-147.221.am zn2023	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250331.(	6.1.131-143.221.am zn2023	18 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central), Ásia-Pacífico (Tóquio)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250303.(	6.1.129-138.220.am zn2023	27 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250218.1	6.1.128-136.201.amzn2023	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderabad), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250211.(	6.1.127-135.201.am zn2023	26 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

Versões Java padrão 7.8.0

O Amazon EMR versão 7.8 e superior vem com o Amazon Corretto 17 (construído no OpenJDK) por padrão para aplicativos que suportam o Corretto 17 (JDK 17), com exceção do Apache Livy.

A tabela a seguir mostra as versões padrão do Java para aplicativos no Amazon EMR 7.8.0. Se você quiser alterar a JVM padrão no cluster, siga as instruções em Configurar aplicações para usar uma Máquina Virtual Java específica para cada aplicação executada no cluster. Você só pode usar uma versão de runtime do Java para um cluster. O Amazon EMR não é compatível com a execução de diferentes nós ou aplicações em diferentes versões de runtime no mesmo cluster.

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
Delta	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Tez	17, 11, 8
Trino	21, 17
Zeppelin	17, 11, 8
Pig	17, 11, 8
Zookeeper	17, 11, 8

Versões do componente 7.8.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *CommunityVersion*-amzn-*EmrVersion*. O *EmrVersion* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado myapp-component com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como 2.2-amzn-2.

Componente	Versão	Descrição
adot-java-agent	1.31.0	Um Java Agent que coleta métricas de daemons de aplicações.
delta	3.3.0-amzn-0	O Delta Lake é um formato de tabela aberto para grandes conjuntos de dados analíticos
emr-amazon-cloudwatch-agent	1.300032.2-amzn-0	Um aplicativo que coleta métricas internas em nível de sistema e métricas de aplicativos personalizadas de instâncias da Amazon. EC2
emr-ddb	5.6.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.17.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.21.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.18.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway

Componente	Versão	Descrição
emr-s3-dist-cp	2.38.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	2.17.0	EMR S3Select Connector
emr-wal-cli	1.3.3	Cli usada para lista/exclusão de emrwal.
emrfs	2.69,0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.20.0-aman-2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.20.0-aman-2	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
hadoop-client	3.4.1-amzn-0	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	3.4.1-amzn-0	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.4.1-amzn-0	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.4.1-amzn-0	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.

Componente	Versão	Descrição
hadoop-hdfs-journalnode	3.4.1-amzn-0	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.4.1-amzn-0	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.4.1-amzn-0	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	3.4.1-amzn-0	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.4.1-amzn-0	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	3.4.1-amzn-0	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	3.4.1-amzn-0	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	2.6.1-amzn-2	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.

Componente	Versão	Descrição
hbase-region-server	2.6.1-amzn-2	Serviço para atender a uma ou mais HBase regiões.
hbase-client	2.6.1-amzn-2	HBase cliente de linha de comando.
hbase-rest-server	2.6.1-amzn-2	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.6.1-amzn-2	Serviço que fornece um endpoint Thrift para. HBase
hbase-operator-tools	2.6.1-amzn-2	Ferramenta de reparo para HBase clusters Apache.
hcatalog-client	3.1.3-aman-17	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.3-aman-17	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	3.1.3-aman-17	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.3-aman-17	O cliente da linha de comando do Hive.
hive-hbase	3.1.3-aman-17	Cliente do Hive-hbase.

Componente	Versão	Descrição
hive-metastore-server	3.1.3-aman-17	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.3-aman-17	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.15.0-amzn-5	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.15.0-amzn-5	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-trino	0.15.0-amzn-5	Biblioteca de pacotes para execução do Presto com o Hudi.
hudi-spark	0.15.0-amzn-5	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.11.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
iceberg	1.7.1-amzn-1	Apache Iceberg é um formato de tabela aberta para conjuntos de dados analíticos imensos
jupyterhub	1.5.0	Servidor de vários usuários para blocos de anotações Jupyter

Componente	Versão	Descrição
livy-server	0.8.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.
nvidia-cuda	12.3.0	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitações de fluxo de trabalho do Oozie.
opencv	4.7.0	Biblioteca de Visão Computacional de Código Aberto.
phoenix-library	5.2.1	As bibliotecas do phoenix para servidor e cliente
phoenix-connectors	6.0.0	Conectores do Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica

Componente	Versão	Descrição
presto-coordinator	0,287-amzn-2	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0,287-amzn-2	O serviço que executa partes de uma consulta.
presto-client	0,287-amzn-2	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
trino-coordinator	467-amzn-1	Serviço que aceita consultas e gerencia a execução de consultas entre trino-workers.
trino-worker	467-amzn-1	O serviço que executa partes de uma consulta.
trino-client	467-amzn-1	Cliente de linha de comando do Trino instalado em mestres de espera de um cluster de alta disponibilidade (HA) em que o servidor do Trino não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.3.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger

Componente	Versão	Descrição
spark-client	3.5.4-amzn-0	Os clientes da linha de comando do Spark.
spark-history-server	3.5.4-amzn-0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.5.4-amzn-0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.5.4-amzn-0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
spark-rapids	24.12.0-aman-0	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs
tensorflow	2.16.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.10.2-amzn-15	O aplicativo e as bibliotecas do YARN para Tez.
tez-on-worker	0.10.2-amzn-15	A aplicação e as bibliotecas do YARN nós de processam ento.
zeppelin-server	0.11.1	O notebook baseado na web que permite um data analytics interativo.

Componente	Versão	Descrição
zookeeper-server	3.9.3-amzn-0	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.9.3-amzn-0	ZooKeeper cliente de linha de comando.

Classificações de configuração 7.8.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

classificações do emr-7.8.0

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Classificações	Descrição	Ações de reconfiguração
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.

Classificações	Descrição	Ações de reconfiguração
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Alterar a página inicial do KMS do Hadoop	Not available.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.

Classificações	Descrição	Ações de reconfiguração
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.	Not available.
iceberg-defaults	Alterar os valores no arquivo iceberg-defaults.conf do Iceberg.	Not available.
delta-defaults	Alterar os valores no arquivo delta-defaults.conf do Delta.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j2	Alterar as configurações de log4j2.properties do Livy.	Restarts Livy Server.

Classificações	Descrição	Ações de reconfiguração
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j2	Altere os valores no arquivo log4j2.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.

Classificações	Descrição	Ações de reconfiguração
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-lakeformation	Alterar os valores no arquivo lakeformation.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
trino-log	Alterar os valores no arquivo log.properties do Trino.	Restarts Trino-Server (for Trino)
trino-config	Alterar os valores no arquivo config.properties do Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Alterar os valores no arquivo password-authenticator.properties do Trino.	Restarts Trino-Server (for Trino)
trino-env	Alterar os valores no arquivo trino-env.sh do Trino.	Restarts Trino-Server (for Trino)
trino-node	Alterar os valores no arquivo node.properties do Trino.	Not available.
trino-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Trino.	Not available.
trino-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Trino.	Not available.
trino-connector-delta	Alterar os valores no arquivo delta.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Alterar os valores no arquivo hive.properties do Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Alterar os valores no arquivo exchange-manager.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Alterar os valores no arquivo iceberg.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Alterar os valores no arquivo hudi.properties do Trino.	Restarts Trino-Server (for Trino)

Classificações	Descrição	Ações de reconfiguração
trino-connector-jmx	Alterar os valores no arquivo jmx.properties do Trino.	Not available.
trino-connector-kafka	Alterar os valores no arquivo kafka.properties do Trino.	Not available.
trino-connector-localfile	Alterar os valores no arquivo localfile.properties do Trino.	Not available.
trino-connector-memory	Alterar os valores no arquivo memory.properties do Trino.	Not available.
trino-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Trino.	Not available.
trino-connector-mysql	Alterar os valores no arquivo mysql.properties do Trino.	Not available.
trino-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Trino.	Not available.
trino-connector-raptor	Alterar os valores no arquivo raptor.properties do Trino.	Not available.
trino-connector-redis	Alterar os valores no arquivo redis.properties do Trino.	Not available.
trino-connector-redshift	Alterar os valores no arquivo redshift.properties do Trino.	Not available.
trino-connector-tpch	Alterar os valores no arquivo tpch.properties do Trino.	Not available.
trino-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Trino.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Alterar os valores no arquivo kms-logback.xml do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j2	Alterar os valores no arquivo log4j2.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie and HiveServer2.

Classificações	Descrição	Ações de reconfiguração
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.
zookeeper-logback	Altere os valores no ZooKeeper arquivo logback.xml do.	Restarts Zookeeper server.
emr-metrics	Altere as configurações de métricas do EMR desse nó.	Restarts the CloudWatchAgent service.

Registro de alterações 7.8.0

Registro de alterações da versão 7.8.0 e das notas de lançamento

Data	Event	Descrição
2025-03-07	Publicação de documentos	Notas de lançamento do Amazon EMR 7.8.0 publicadas pela primeira vez
2025-03-06	Implantação concluída	Amazon EMR 7.8.0 totalmente implantado em todas as regiões suportadas
2025-02-28	Lançamento inicial limitado	Amazon EMR 7.8.0 foi implantado pela primeira vez nas regiões comerciais iniciais

Amazon EMR versão 7.7.0

Ciclo de vida compatível com 7.7.0

A tabela a seguir descreve as datas de ciclo de vida suportadas para a versão 7.7.0 do Amazon EMR.

Fase de suporte	Data
Data da versão inicial	6 de fevereiro de 2025
Suporte padrão até	5 de fevereiro de 2027
Término do suporte	6 de fevereiro de 2027
Fim da vida útil	6 de fevereiro de 2028

Componentes da versão 7.7.0

O Amazon EMR 7.7.0 inclui os seguintes aplicativos:

Mecanismos centrais:

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

Extras:

Hue, Oozie JupyterHub, Porco, Presto, Zeppelin TensorFlow

Para obter mais detalhes sobre os pacotes incluídos, consulte Versões do aplicativo nas versões 7.x do Amazon EMR. Para obter mais informações sobre os mecanismos centrais e extras, consulte Amazon EMR Standard Support.

Versões do aplicativo 7.7.0

Essa versão inclui os seguintes aplicativos: AmazonCloudWatchAgentDeltaFlink, HBase, HCatalogHadoop, Hive, Hudi, Hue, IcebergJupyterEnterpriseGateway, JupyterHub, Livy, Oozie, PhoenixPig, Presto, Spark, TensorFlow, TezTrino, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-7.7.0	emr-7.6.0	emr-7.5.0	emr-7.4.0
AWS SDK for Java	2.29.44, 1.12.780	2.29.25, 1.12.779	2.28.8, 1.12.772	2.28.8, 1.12.772
Python	3,9, 3,11	3,9, 3,11	3,9, 3,11	3,9, 3,11
Scala	2.12.18	2.12.18	2.12.18	2.12.18

	emr-7.7.0	emr-7.6.0	emr-7.5.0	emr-7.4.0
AmazonCloudWatchAgent	1.300032.2	1.300032.2	1.300032.2	1.300032.2
Delta	3.2.1-amzn-2	3.2.1-amzn-1	3.2.1-amzn-0	3.2.0-amzn-1
Flink	1.20.0	1.20.0	1.19.1	1.19.1
Ganglia	-	-	-	-
HBase	2.6.1	2.6.1	2.5.10	2.5.5
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.4.0	3.4.0	3.4.0	3.4.0
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-aman-4	0.15.0-aman-3	0.15.0-amzn-2	0.15.0-amzn-1
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.7.1-amzn-0	1.6.1-amzn-2	1.6.1-amzn-1	1.6.1-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.8.0
MXNet	-	-	-	-
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.2.1	5.2.0	5.2.0	5.2.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0

	emr-7.7.0	emr-7.6.0	emr-7.5.0	emr-7.4.0
Presto	0,287	0,287	0,287	0,287
Spark	3.5.3	3.5.3	3.5.2	3.5.2
Sqoop	-	-	-	1.4.7
TensorFlow	2.16.1	2.16.1	2.16.1	2.16.1
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	467	457	446	446
Zeppelin	0.11.1	0.11.1	0.11.1	0.11.1
ZooKeeper	3.9.2	3.9.2	3.9.2	3.9.2

Detalhes dos componentes da versão 7.7.0

Consulte a tabela a seguir para obter mais informações sobre os pacotes extras no Amazon EMR 7.7.0.

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
AmazonCloudWatchAgent	Serviços	hadoop-hdfs-journalnode	Extras	3.4.0-amzn-3-1.amzn2023
AmazonCloudWatchAgent	Serviços	ranger-kms	Extras	2.0.0-1.amzn2023
AmazonCloudWatchAgent	Serviços	s3-dist-cp	Extras	2.37.0-1.amzn2023
AmazonCloudWatchAgent	Serviços	zookeeper	Extras	3.9.2.amzn.1-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
AmazonCloudWatchAgent	Serviços	zookeeper-server	Extras	3.9.2.amzn.1-1.amzn2023
Flink	Serviços	hadoop-client	Extras	3.4.0.amzn.3-1.amzn2023
Flink	Serviços	hadoop-hdfs	Extras	3.4.0.amzn.3-1.amzn2023
Flink	Serviços	hadoop-hdfs-datanode	Extras	3.4.0.amzn.3-1.amzn2023
Flink	Serviços	hadoop-hdfs-journalnode	Extras	3.4.0.amzn.3-1.amzn2023
Flink	Serviços	hadoop-hdfs-namenode	Extras	3.4.0.amzn.3-1.amzn2023
Flink	Serviços	hadoop-httpfs	Extras	3.4.0.amzn.3-1.amzn2023
Flink	Serviços	hadoop-kms	Extras	3.4.0.amzn.3-1.amzn2023
Flink	Serviços	hadoop-mapreduce	Extras	3.4.0.amzn.3-1.amzn2023
Flink	Serviços	hadoop-mapreduce-historyserver	Extras	3.4.0.amzn.3-1.amzn2023
Flink	Serviços	hadoop-yarn	Extras	3.4.0.amzn.3-1.amzn2023
Flink	Serviços	hadoop-yarn-nodemanager	Extras	3.4.0.amzn.3-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Flink	Serviços	hadoop-yarn-proxyserver	Extras	3.4.0.amzn.3-1.amzn2023
Flink	Serviços	hadoop-yarn-resourcemanager	Extras	3.4.0.amzn.3-1.amzn2023
Flink	Serviços	hadoop-yarn-timelineserver	Extras	3.4.0.amzn.3-1.amzn2023
Flink	Serviços	hudi	Extras	0.15.0.amzn.4-1.amzn2023
Flink	Serviços	ranger-kms	Extras	2.0.0-1.amzn2023
Flink	Serviços	hadoop	Extras	3.4.0.amzn.3-1.amzn2023
HBase	Serviços	s3-dist-cp	Extras	2.37.0-1.amzn2023
HCatalog	Serviços	hadoop-yarn-proxyserver	Extras	3.4.0.amzn.3-1.amzn2023
HCatalog	Serviços	mariadb-connector-java	Extras	
Hive	Serviços	hadoop-httpfs	Extras	3.4.0.amzn.3-1.amzn2023
Hive	Serviços	hadoop-yarn-proxyserver	Extras	3.4.0.amzn.3-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Hive	Serviços	hive-hbase	Extras	3.1.3.amzn.16-1.amzn2023
Hive	Serviços	mariadb-connector-java	Extras	2.7.2-1
Hive	Serviços	s3-dist-cp	Extras	2.37.0-1.amzn2023
Phoenix	Serviços	s3-dist-cp	Extras	2.37.0-1.amzn2023
Spark	Serviços	spark-datanucleus	Extras	3.5.3.amzn.1-1.amzn2023
Tez	Serviços	hadoop-mapreduce	Extras	3.4.0.amzn.3-1.amzn2023
Tez	Serviços	hadoop-mapreduce-historyserver	Extras	3.4.0.amzn.3-1.amzn2023
Tez	Serviços	hadoop-yarn-proxyserver	Extras	3.4.0.amzn.3-1.amzn2023
Trino	Serviços	bigtop-utils	Extras	1.2.0-1.amzn2023
Trino	Serviços	hadoop-client	Extras	3.4.0.amzn.3-1.amzn2023
Trino	Serviços	hadoop-hdfs	Extras	3.4.0.amzn.3-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Trino	Serviços	hadoop-hdfs-datanode	Extras	3.4.0.amzn.3-1.amzn2023
Trino	Serviços	hadoop-hdfs-journalnode	Extras	3.4.0.amzn.3-1.amzn2023
Trino	Serviços	hadoop-hdfs-namenode	Extras	3.4.0.amzn.3-1.amzn2023
Trino	Serviços	hadoop-kms	Extras	3.4.0.amzn.3-1.amzn2023
Trino	Serviços	hadoop-mapreduce	Extras	3.4.0.amzn.3-1.amzn2023
Trino	Serviços	hadoop-mapreduce-historyserver	Extras	3.4.0.amzn.3-1.amzn2023
Trino	Serviços	hadoop-yarn	Extras	3.4.0.amzn.3-1.amzn2023
Trino	Serviços	hadoop-yarn-nodemanager	Extras	3.4.0.amzn.3-1.amzn2023
Trino	Serviços	hadoop-yarn-proxyserver	Extras	3.4.0.amzn.3-1.amzn2023
Trino	Serviços	hadoop-yarn-resourcemanager	Extras	3.4.0.amzn.3-1.amzn2023
Trino	Serviços	hadoop-yarn-timelineserver	Extras	3.4.0.amzn.3-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Trino	Serviços	hive	Extras	3.1.3.amzn.16-1.amzn2023
Trino	Serviços	hive-hcatalog-server	Extras	3.1.3.amzn.16-1.amzn2023
Trino	Serviços	mariadb-connector-java	Extras	2.7.2-1
Trino	Serviços	ranger-kms	Extras	2.0.0-1.amzn2023
Trino	Serviços	zookeeper	Extras	3.9.2.amzn.1-1.amzn2023
Trino	Serviços	zookeeper-server	Extras	3.9.2.amzn.1-1.amzn2023
Trino	Serviços	bigtop-groovy	Extras	2.5.4-1.amzn2023
Trino	Serviços	bigtop-jsvc	Extras	1.2.4-1.amzn2023
Trino	Serviços	hadoop	Extras	3.4.0.amzn.3-1.amzn2023
Trino	Serviços	hadoop-lzo	Extras	0.4.19-1.amzn2023
Trino	Serviços	hive-hcatalog	Extras	3.1.3.amzn.16-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Trino	Serviços	hive-jdbc	Extras	3.1.3.amzn.16-1.amzn2023
Zookeeper	Serviços	emrfs	Extras	3.9.2.amzn.1-1.amzn2023
Zookeeper	Serviços	hadoop-client	Extras	3.9.2.amzn.1-1.amzn2023
Zookeeper	Serviços	hadoop-hdfs	Extras	3.9.2.amzn.1-1.amzn2023
Zookeeper	Serviços	hadoop-hdfs-datanode	Extras	3.9.2.amzn.1-1.amzn2023
Zookeeper	Serviços	hadoop-hdfs-journalnode	Extras	3.9.2.amzn.1-1.amzn2023
Zookeeper	Serviços	hadoop-hdfs-namenode	Extras	3.9.2.amzn.1-1.amzn2023
Zookeeper	Serviços	hadoop-httpfs	Extras	3.9.2.amzn.1-1.amzn2023
Zookeeper	Serviços	hadoop-kms	Extras	3.9.2.amzn.1-1.amzn2023
Zookeeper	Serviços	hadoop-mapreduce	Extras	3.4.0.amzn.3-1.amzn2023
Zookeeper	Serviços	hadoop-mapreduce-historyserver	Extras	3.4.0.amzn.3-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Zookeeper	Serviços	hadoop-yarn	Extras	3.4.0.amzn.3-1.amzn2023
Zookeeper	Serviços	hadoop-yarn-nodemanager	Extras	3.4.0.amzn.3-1.amzn2023
Zookeeper	Serviços	hadoop-yarn-proxyserver	Extras	3.4.0.amzn.3-1.amzn2023
Zookeeper	Serviços	hadoop-yarn-resourcemanager	Extras	3.4.0.amzn.3-1.amzn2023
Zookeeper	Serviços	hadoop-yarn-timelineserver	Extras	3.4.0.amzn.3-1.amzn2023
Zookeeper	Serviços	ranger-kms	Extras	2.0.0-1.amzn2023

Notas de lançamento da versão 7.7.0

As notas de lançamento a seguir incluem informações sobre a versão 7.7.0 do Amazon EMR.

Novos recursos

- Atualizações da aplicação: Amazon EMR 7.7.0 application upgrades include Delta 3.2.1-amzn-2, Hudi 0.15.0-amzn-4, Iceberg 1.7.1-amzn-0, and Trino 467.

Problemas conhecidos

- Não aplicável a esta versão.

Alterações, melhorias e problemas resolvidos

- A versão Iceberg em uso a partir do EMR 7.7.0 não oferece mais suporte ao Java 8. Além disso, o Iceberg está excluído das seguintes imagens do Java 8: `emr-7.7.0-java8-latest` e. `emr-7.7.0-spark-rapids-java8-latest`
- Lake Formation para FGAC com Amazon EMR no EKS — Com o Amazon EMR versão 7.7 e superior, você pode aproveitar o AWS Lake Formation para aplicar controles de acesso refinados em tabelas do catálogo de dados que são apoiadas pelo Amazon S3. Esse recurso permite que você configure controles de acesso em nível de tabela, linha, coluna e célula para consultas de leitura em seu Amazon Amazon EMR no EKS Spark Jobs. Para obter mais informações, consulte Usando o Amazon Amazon EMR no EKS com o AWS Lake Formation para um controle de acesso refinado.
-

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250527.1	6.1.134-152.225.amzn2023	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein), China (Pequim), Ásia-Pacífico (Malásia), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Leste dos EUA (Norte da Virgínia), Ásia-Pací

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Malásia), Europa (Londres), Ásia-Pacífico (Melbourne), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (AWS GovCloud AWS GovCloud Oregon), México (Central) , Oeste dos EUA (Norte da Califórnia), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv), Ásia-Pacífico (Taipei), Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250512.(	6.1.134-152.225.am zn2023	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Hyderabad), Ásia-Pacífico (Jacarta)), Ásia-Pacífico (Malásia), Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Oriente Médio (Oriente Médio) (Oriente Médio (Oriente Médio)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Emirados Árabes Unidos), Canadá (Central) , Oeste do Canadá (Calgary), (Leste dos EUA), AWS GovCloud (AWS GovCloud Oeste dos EUA), China (Pequim), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250428.1	6.1.134-150.224.amzn2023	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Jacarta) Arta), Ásia-Pacífico (Malásia), Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Malásia) , Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (EAU), Canadá (Central) , Oeste do Canadá

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Calgary), (Leste dos EUA), AWS GovCloud (Oeste dos EUA), China (AWS GovCloud Pequim), China (Ningxia), Israel (Tel Aviv), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250414.(	6.1.132-147.221.am zn2023	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250331.(	6.1.131-143.221.am zn2023	18 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central), Ásia-Pacífico (Tóquio)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250303.(	6.1.129-138.220.am zn2023	27 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250218.1	6.1.128-136.201.amzn2023	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderabad), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250211.(	6.1.127-135.201.am zn2023	26 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

7.7.0 versões padrão do Java

O Amazon EMR versão 7.7 e superior vem com o Amazon Corretto 17 (construído no OpenJDK) por padrão para aplicativos que suportam o Corretto 17 (JDK 17), com exceção do Apache Livy.

A tabela a seguir mostra as versões padrão do Java para aplicativos no Amazon EMR 7.7.0. Se você quiser alterar a JVM padrão no cluster, siga as instruções em Configurar aplicações para usar uma Máquina Virtual Java específica para cada aplicação executada no cluster. Você só pode usar uma versão de runtime do Java para um cluster. O Amazon EMR não é compatível com a execução de diferentes nós ou aplicações em diferentes versões de runtime no mesmo cluster.

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
Delta	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Tez	17, 11, 8
Trino	21, 17
Zeppelin	17, 11, 8
Pig	17, 11, 8
Zookeeper	17, 11, 8

Versões de componentes 7.7.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *CommunityVersion*-amzn-*EmrVersion*. O *EmrVersion* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado myapp-component com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como 2.2-amzn-2.

Componente	Versão	Descrição
adot-java-agent	1.31.0	Um Java Agent que coleta métricas de daemons de aplicações.
delta	3.2.1-amzn-2	O Delta Lake é um formato de tabela aberto para grandes conjuntos de dados analíticos
emr-amazon-cloudwatch-agent	1.300032.2-amzn-0	Um aplicativo que coleta métricas internas em nível de sistema e métricas de aplicativos personalizadas de instâncias da Amazon. EC2
emr-ddb	5.6.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.16.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.20.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.18.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway

Componente	Versão	Descrição
emr-s3-dist-cp	2.37.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	2.16.0	EMR S3Select Connector
emr-wal-cli	1.3.2	Cli usada para lista/exclusão de emrwal.
emrfs	2.68,0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.20.0-amzn-1	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.20.0-amzn-1	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
hadoop-client	3.4.0-aman-3	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	3.4.0-aman-3	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.4.0-aman-3	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.4.0-aman-3	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.

Componente	Versão	Descrição
hadoop-hdfs-journalnode	3.4.0-aman-3	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.4.0-aman-3	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.4.0-aman-3	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	3.4.0-aman-3	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.4.0-aman-3	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	3.4.0-aman-3	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	3.4.0-aman-3	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	2.6.1-amzn-1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.

Componente	Versão	Descrição
hbase-region-server	2.6.1-amzn-1	Serviço para atender a uma ou mais HBase regiões.
hbase-client	2.6.1-amzn-1	HBase cliente de linha de comando.
hbase-rest-server	2.6.1-amzn-1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.6.1-amzn-1	Serviço que fornece um endpoint Thrift para. HBase
hbase-operator-tools	2.6.1-amzn-1	Ferramenta de reparo para HBase clusters Apache.
hcatalog-client	3.1.3-aman-16	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.3-aman-16	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	3.1.3-aman-16	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.3-aman-16	O cliente da linha de comando do Hive.
hive-hbase	3.1.3-aman-16	Cliente do Hive-hbase.

Componente	Versão	Descrição
hive-metastore-server	3.1.3-aman-16	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.3-aman-16	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.15.0-aman-4	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.15.0-aman-4	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-trino	0.15.0-aman-4	Biblioteca de pacotes para execução do Presto com o Hudi.
hudi-spark	0.15.0-aman-4	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.11.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
iceberg	1.7.1-amzn-0	Apache Iceberg é um formato de tabela aberta para conjuntos de dados analíticos imensos
jupyterhub	1.5.0	Servidor de vários usuários para blocos de anotações Jupyter

Componente	Versão	Descrição
livy-server	0.8.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.
nvidia-cuda	12.3.0	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitações de fluxo de trabalho do Oozie.
opencv	4.7.0	Biblioteca de Visão Computacional de Código Aberto.
phoenix-library	5.2.1	As bibliotecas do phoenix para servidor e cliente
phoenix-connectors	6.0.0	Conectores do Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica

Componente	Versão	Descrição
presto-coordinator	0,287-amzn-1	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0,287-amzn-1	O serviço que executa partes de uma consulta.
presto-client	0,287-amzn-1	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
trino-coordinator	467-amzn-0	Serviço que aceita consultas e gerencia a execução de consultas entre trino-workers.
trino-worker	467-amzn-0	O serviço que executa partes de uma consulta.
trino-client	467-amzn-0	Cliente de linha de comando do Trino instalado em mestres de espera de um cluster de alta disponibilidade (HA) em que o servidor do Trino não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.3.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger

Componente	Versão	Descrição
spark-client	3.5.3-amzn-1	Os clientes da linha de comando do Spark.
spark-history-server	3.5.3-amzn-1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.5.3-amzn-1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.5.3-amzn-1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
spark-rapids	24.10.1-amzn-1	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs
tensorflow	2.16.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.10.2-amzn-14	O aplicativo e as bibliotecas do YARN para Tez.
tez-on-worker	0.10.2-amzn-14	A aplicação e as bibliotecas do YARN nós de processam ento.
zeppelin-server	0.11.1	O notebook baseado na web que permite um data analytics interativo.

Componente	Versão	Descrição
zookeeper-server	3.9.2-amzn-1	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.9.2-amzn-1	ZooKeeper cliente de linha de comando.

Classificações de configuração 7.7.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

classificações do emr-7.7.0

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Classificações	Descrição	Ações de reconfiguração
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.

Classificações	Descrição	Ações de reconfiguração
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Alterar a página inicial do KMS do Hadoop	Not available.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.

Classificações	Descrição	Ações de reconfiguração
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.	Not available.
iceberg-defaults	Alterar os valores no arquivo iceberg-defaults.conf do Iceberg.	Not available.
delta-defaults	Alterar os valores no arquivo delta-defaults.conf do Delta.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j2	Alterar as configurações de log4j2.properties do Livy.	Restarts Livy Server.

Classificações	Descrição	Ações de reconfiguração
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j2	Altere os valores no arquivo log4j2.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.

Classificações	Descrição	Ações de reconfiguração
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-lakeformation	Alterar os valores no arquivo lakeformation.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
trino-log	Alterar os valores no arquivo log.properties do Trino.	Restarts Trino-Server (for Trino)
trino-config	Alterar os valores no arquivo config.properties do Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Alterar os valores no arquivo password-authenticator.properties do Trino.	Restarts Trino-Server (for Trino)
trino-env	Alterar os valores no arquivo trino-env.sh do Trino.	Restarts Trino-Server (for Trino)
trino-node	Alterar os valores no arquivo node.properties do Trino.	Not available.
trino-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Trino.	Not available.
trino-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Trino.	Not available.
trino-connector-delta	Alterar os valores no arquivo delta.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Alterar os valores no arquivo hive.properties do Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Alterar os valores no arquivo exchange-manager.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Alterar os valores no arquivo iceberg.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Alterar os valores no arquivo hudi.properties do Trino.	Restarts Trino-Server (for Trino)

Classificações	Descrição	Ações de reconfiguração
trino-connector-jmx	Alterar os valores no arquivo jmx.properties do Trino.	Not available.
trino-connector-kafka	Alterar os valores no arquivo kafka.properties do Trino.	Not available.
trino-connector-localfile	Alterar os valores no arquivo localfile.properties do Trino.	Not available.
trino-connector-memory	Alterar os valores no arquivo memory.properties do Trino.	Not available.
trino-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Trino.	Not available.
trino-connector-mysql	Alterar os valores no arquivo mysql.properties do Trino.	Not available.
trino-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Trino.	Not available.
trino-connector-raptor	Alterar os valores no arquivo raptor.properties do Trino.	Not available.
trino-connector-redis	Alterar os valores no arquivo redis.properties do Trino.	Not available.
trino-connector-redshift	Alterar os valores no arquivo redshift.properties do Trino.	Not available.
trino-connector-tpch	Alterar os valores no arquivo tpch.properties do Trino.	Not available.
trino-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Trino.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Alterar os valores no arquivo kms-logback.xml do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j2	Alterar os valores no arquivo log4j2.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie and HiveServer2.

Classificações	Descrição	Ações de reconfiguração
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.
zookeeper-logback	Altere os valores no ZooKeeper arquivo logback.xml do.	Restarts Zookeeper server.
emr-metrics	Altere as configurações de métricas do EMR desse nó.	Restarts the CloudWatchAgent service.

Registro de alterações 7.7.0

Registro de alterações da versão 7.7.0 e das notas de lançamento

Data	Event	Descrição
2025-02-07	Publicação de documentos	Notas de lançamento do Amazon EMR 7.7.0 publicadas pela primeira vez
2025-02-06	Implantação concluída	Amazon EMR 7.7.0 totalmente implantado em todas as regiões suportadas
2025-01-30	Lançamento inicial limitado	Amazon EMR 7.7.0 foi implantado pela primeira vez nas regiões comerciais iniciais

Amazon EMR versão 7.6.0

Ciclo de vida compatível com 7.6.0

A tabela a seguir descreve as datas de ciclo de vida suportadas para a versão 7.6.0 do Amazon EMR.

Fase de suporte	Data
Data da versão inicial	10 de janeiro de 2025
Suporte padrão até	09 de janeiro de 2027
Término do suporte	10 de janeiro de 2027
Fim da vida útil	10 de janeiro de 2028

Componentes da versão 7.6.0

O Amazon EMR 7.6.0 inclui os seguintes aplicativos:

Mecanismos centrais:

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

Extras:

Hue, Oozie JupyterHub, Porco, Presto, Zeppelin TensorFlow

Para obter mais detalhes sobre os pacotes incluídos, consulte Versões do aplicativo nas versões 7.x do Amazon EMR. Para obter mais informações sobre os mecanismos centrais e extras, consulte Amazon EMR Standard Support.

Versões do aplicativo 7.6.0

Essa versão inclui os seguintes aplicativos: AmazonCloudWatchAgent, Delta, Flink, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, Oozie, Phoenix, Pig, Presto, Spark, TensorFlow, TezTrino, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-7.6.0	emr-7.5.0	emr-7.4.0	emr-7.3.0
AWS SDK for Java	2.29.25, 1.12.779	2.28.8, 1.12.772	2.28.8, 1.12.772	2.25.70, 1.12.747
Python	3,9, 3,11	3,9, 3,11	3,9, 3,11	3,9, 3,11
Scala	2.12.18	2.12.18	2.12.18	2.12.18

	emr-7.6.0	emr-7.5.0	emr-7.4.0	emr-7.3.0
AmazonCloudWatchAgent	1.300032.2	1.300032.2	1.300032.2	1.300032.2
Delta	3.2.1-amzn-1	3.2.1-amzn-0	3.2.0-amzn-1	3.2.0-amzn-0
Flink	1.20.0	1.19.1	1.19.1	1.18.1
Ganglia	-	-	-	-
HBase	2.6.1	2.5.10	2.5.5	2.4.17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.4.0	3.4.0	3.4.0	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-aman-3	0.15.0-amzn-2	0.15.0-amzn-1	0.15.0-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.6.1-amzn-2	1.6.1-amzn-1	1.6.1-amzn-0	1.5.2-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.8.0
MXNet	-	-	-	-
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.2.0	5.2.0	5.2.0	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0

	emr-7.6.0	emr-7.5.0	emr-7.4.0	emr-7.3.0
Presto	0,287	0,287	0,287	0,285
Spark	3.5.3	3.5.2	3.5.2	3.5.1
Sqoop	-	-	1.4.7	1.4.7
TensorFlow	2.16.1	2.16.1	2.16.1	2.16.1
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	457	446	446	442
Zeppelin	0.11.1	0.11.1	0.11.1	0.11.1
ZooKeeper	3.9.2	3.9.2	3.9.2	3.9.1

7.6.0 vulnerabilidades e exposições comuns

A tabela a seguir lista todos os CVEs que não afetam os clusters do EMR que são executados nas configurações recomendadas do Amazon EMR 7.6.0. O Amazon EMR depende do código upstream aberto para disponibilidade de correções e fornecerá a versão estável mais recente como parte dos componentes do mecanismo central em até 90 dias após a verificação das correções pelo Amazon EMR.

Essa tabela foi atualizada em 10 de janeiro de 2025.

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-15250	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-15250
CVE-2020-17521	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-17 521
CVE-2024-12798	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-12798

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2024-12801	BAIXO	https://nvd.nist. gov/vuln/detail/CVE-2024-12801
CVE-2024-47554	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2024-47 554
CVE-2022-1471	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-1471
CVE-2022-25857	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-25 857
CVE-2022-38749	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2022-38749
CVE-2022-38750	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2022-38750
CVE-2022-38751	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2022-38751
CVE-2022-38752	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2022-38752
CVE-2022-41854	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2022-4 1854
CVE-2022-42003	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2003
CVE-2022-42004	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2004
CVE-2023-6378	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-6378
CVE-2023-6481	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-6481

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2024-36114	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2024-3614
CVE-2015-1832	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2015-1832
CVE-2018-1313	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2018-13 13
CVE-2019-10219	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2019-10 219
CVE-2020-10693	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-10 693
CVE-2020-13949	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2020-13955	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-1395
CVE-2020-29582	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-29582
CVE-2020-8908	BAIXO	https://nvd.nist. gov/vuln/detail/CVE-2020-8908
CVE-2021-28170	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2021-28 170
CVE-2021-34538	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-34 538
CVE-2021-38153	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2021-38 153
CVE-2022-24329	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2022-24 329

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2022-25647	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-25 647
CVE-2022-36364	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3636 4
CVE-2022-39135	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-39135
CVE-2022-42889	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-4289
CVE-2022-46337	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-46337
CVE-2022-46751	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-46751
CVE-2023-1370	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2023-20863	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-20863
CVE-2023-20883	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2083
CVE-2023-25194	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-25194
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-33201	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-33201
CVE-2023-33202	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-33202

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2023-34453	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3445 3
CVE-2023-34454	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3445
CVE-2023-34455	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3445
CVE-2023-3635	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3635
CVE-2023-39410	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-42503	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-42503
CVE-2023-43642	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2023-5072	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-5072
CVE-2023-51074	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-510 74
CVE-2024-23945	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2024-23945
CVE-2024-25710	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-25710
CVE-2024-26308	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-26308
CVE-2024-29857	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2024-29 857

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2024-30171	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2024-30 171
CVE-2024-31141	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2024 - 31 141
CVE-2024-38808	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2024-38808
CVE-2024-38820	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-38820
CVE-2024-47561	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-47 56
CVE-2024-52046	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-5 2046
CVE-2024-53990	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-53990
CVE-2024-56128	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2024-56128
CVE-2024-7254	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-7254
GHSA-CHFM-68VV-PvW5	BAIXO	Nenhum link disponível.
CVE-2024-47535	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2024-47 535
CVE-2012-5783	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2012-5783
CVE-2012-6153	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2012-6153

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2016-5018	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2016-5018
CVE-2016-6796	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2016-6796
CVE-2018-10237	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2018-10 237
CVE-2018-11798	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2018-1 1798
CVE-2019-0205	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-0205
CVE-2020-15522	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-1552
CVE-2020-1945	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-1945
CVE-2021-36373	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2021-36373
CVE-2021-36374	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2021-36374
CVE-2018-1000180	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-1000180
CVE-2020-26939	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-26939
CVE-2020-28052	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-28052
CVE-2021-29425	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2021-29425

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2015-5237	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2015-5237
CVE-2017-15095	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2017-15095
CVE-2017-17485	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2017-17485
CVE-2017-7525	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2017-7525
CVE-2018-11307	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-11 307
CVE-2018-12022	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2018-12022
CVE-2018-12023	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-12 023
CVE-2018-14718	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14 718
CVE-2018-14719	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14 719
CVE-2018-14720	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14 720
CVE-2018-14721	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14 721
CVE-2018-19360	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-19360
CVE-2018-19361	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-1936 1

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2018-19362	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-1936 2
CVE-2018-5968	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-5968
CVE-2018-7489	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-7 489
CVE-2019-12086	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-12 2086
CVE-2019-12384	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2019-12 384
CVE-2019-12814	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2019-12814
CVE-2019-14379	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14 379
CVE-2019-14439	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-14 439
CVE-2019-14540	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14 540
CVE-2019-14892	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14 892
CVE-2019-14893	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14 893
CVE-2019-16335	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16 335
CVE-2019-16942	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16 942

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2019-16943	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16 943
CVE-2019-17267	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17 267
CVE-2019-17531	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17 531
CVE-2019-20330	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20330
CVE-2020-10650	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10650
CVE-2020-10672	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10 672
CVE-2020-10673	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10 673
CVE-2020-10968	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10 968
CVE-2020-10969	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10 969
CVE-2020-11111	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-1111
CVE-2020-11112	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-1112
CVE-2020-11113	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-11-13
CVE-2020-11619	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-11619

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-11620	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-11620
CVE-2020-14060	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14060
CVE-2020-14061	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14061
CVE-2020-14062	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14062
CVE-2020-14195	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14 195
CVE-2020-24616	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-24 616
CVE-2020-24750	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-24750
CVE-2020-35490	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-35490
CVE-2020-35491	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-35491
CVE-2020-35728	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-35728
CVE-2020-36179	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36179
CVE-2020-36180	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36180
CVE-2020-36181	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36181

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-36182	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36182
CVE-2020-36183	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36183
CVE-2020-36184	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36184
CVE-2020-36185	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36185
CVE-2020-36186	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36186
CVE-2020-36187	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36187
CVE-2020-36188	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-3618
CVE-2020-36189	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36189
CVE-2020-36518	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36518
CVE-2020-8840	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-8840
CVE-2020-9546	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9546
CVE-2020-9547	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9547
CVE-2020-9548	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9548

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2021-20190	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-2019 0
CVE-2021-22569	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2021-22569
CVE-2021-22570	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2021-22570
CVE-2022-3171	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3510
CVE-2023-50387	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-50387
CVE-2023-50868	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2023-50868
CVE-2023-52428	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-52428
CVE-2024-23454	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2024-23454
CVE-2024-23944	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2024-2394
CVE-2024-25638	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2024-25 638
CVE-2024-29025	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2024-29 025

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2024-29131	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2024-29 131
CVE-2024-29133	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2024-2913
CVE-2024-51504	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2024-51504
CVE-2024-6763	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-6763
CVE-2024-8184	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-8184
CVE-2022-40149	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40149
CVE-2022-40150	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40150
CVE-2022-45685	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-45 685
CVE-2022-45693	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-45693
CVE-2023-1436	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2013-7285	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2013-7285
CVE-2014-8125	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2014-8 125
CVE-2016-3674	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2016-3674

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2017-7957	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2017-79 57
CVE-2019-10173	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-10 173
CVE-2020-26217	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-26 217
CVE-2020-26258	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-26258
CVE-2021-21341	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-21341
CVE-2021-21342	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-21342
CVE-2021-21343	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-213 43
CVE-2021-21344	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-213 44
CVE-2021-21345	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-21345
CVE-2021-21346	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-21346
CVE-2021-21347	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-21347
CVE-2021-21348	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-21348
CVE-2021-21349	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-21349

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2021-21350	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-21 350
CVE-2021-21351	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-21 351
CVE-2021-29505	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-29505
CVE-2021-39139	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-39139
CVE-2021-39141	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-39141
CVE-2021-39144	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-3914
CVE-2021-39145	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-39145
CVE-2021-39146	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-39146
CVE-2021-39147	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-39147
CVE-2021-39148	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-39148
CVE-2021-39149	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-39149
CVE-2021-39150	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-39150
CVE-2021-39151	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-39 151

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2021-39152	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-39152
CVE-2021-39153	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-39153
CVE-2021-39154	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-39154
CVE-2021-41411	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-414 11
CVE-2021-43859	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-438 59
CVE-2022-1415	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-1415
CVE-2022-40151	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40151
CVE-2022-41966	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-4 1966
CVE-2024-47072	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2024-470 72
CVE-2019-10172	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-10 172
CVE-2019-10202	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-10 202
CVE-2022-45688	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-4568
CVE-2023-0833	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-0833

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2024-21634	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-21634
CVE-2024-9823	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2024-98 23
CVE-2022-37866	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3786
CVE-2012-0881	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2012-08 81
CVE-2022-23437	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2022-23 437
CVE-2020-13936	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-13936
CVE-2022-37865	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-37865
CVE-2009-2625	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2009-2625
CVE-2017-10355	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2017-1035
CVE-2021-31684	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-3 684
CVE-2020-14338	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-14 338
CVE-2018-2799	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2018-27 99
CVE-2021-27568	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2021-27568

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2013-4002	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2013-4002
CVE-2020-13956	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-13956
CVE-2020-25649	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-25 649
CVE-2023-26049	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-26049
CVE-2024-43382	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2024-43 382
CVE-2023-34462	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-3446
CVE-2021-21290	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2021-21290
CVE-2023-44487	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-44487
CVE-2019-20445	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20 445
GHSA-XPW8-RCWV-8F8p	HIGH (ALTO)	Nenhum link disponível.
CVE-2021-37136	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-37136
CVE-2021-21295	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2021-21295
GHSA-58QW-P7QM-5RVh	BAIXO	Nenhum link disponível.

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2021-37533	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2021-3753
CVE-2019-20444	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20 44
CVE-2022-40152	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40152

Detalhes dos componentes da versão 7.6.0

Consulte a tabela a seguir para obter mais informações sobre os pacotes extras no Amazon EMR 7.6.0.

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
AmazonCloudWatchAgent	Serviços	hadoop-hdfs-journalnode	Extras	3.4.0.amzn.2-1.amzn2023
AmazonCloudWatchAgent	Serviços	ranger-kms	Extras	2.0.0-1.amzn2023
AmazonCloudWatchAgent	Serviços	s3-dist-cp	Extras	2.36.0-1.amzn2023
AmazonCloudWatchAgent	Serviços	zookeeper	Extras	3.9.2.amzn.0-1.amzn2023
AmazonCloudWatchAgent	Serviços	zookeeper-server	Extras	3.9.2.amzn.0-1.amzn2023
Flink	Serviços	hadoop-client	Extras	3.4.0.amzn.2-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Flink	Serviços	hadoop-hdfs	Extras	3.4.0.amzn.2-1.amzn2023
Flink	Serviços	hadoop-hdfs-datanode	Extras	3.4.0.amzn.2-1.amzn2023
Flink	Serviços	hadoop-hdfs-journalnode	Extras	3.4.0.amzn.2-1.amzn2023
Flink	Serviços	hadoop-hdfs-namenode	Extras	3.4.0.amzn.2-1.amzn2023
Flink	Serviços	hadoop-httpfs	Extras	3.4.0.amzn.2-1.amzn2023
Flink	Serviços	hadoop-kms	Extras	3.4.0.amzn.2-1.amzn2023
Flink	Serviços	hadoop-mapreduce	Extras	3.4.0.amzn.2-1.amzn2023
Flink	Serviços	hadoop-mapreduce-historyserver	Extras	3.4.0.amzn.2-1.amzn2023
Flink	Serviços	hadoop-yarn	Extras	3.4.0.amzn.2-1.amzn2023
Flink	Serviços	hadoop-yarn-nodemanager	Extras	3.4.0.amzn.2-1.amzn2023
Flink	Serviços	hadoop-yarn-proxyserver	Extras	3.4.0.amzn.2-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Flink	Serviços	hadoop-yarn-resourcemanager	Extras	3.4.0.amzn.2-1.amzn2023
Flink	Serviços	hadoop-yarn-timelineserver	Extras	3.4.0.amzn.2-1.amzn2023
Flink	Serviços	hudi	Extras	0.15.0.amzn.3-1.amzn2023
Flink	Serviços	ranger-kms	Extras	2.0.0-1.amzn2023
Flink	Serviços	hadoop	Extras	3.4.0.amzn.2-1.amzn2023
HBase	Serviços	s3-dist-cp	Extras	2.36.0-1.amzn2023
HCatalog	Serviços	hadoop-yarn-proxyserver	Extras	3.4.0.amzn.2-1.amzn2023
HCatalog	Serviços	mariadb-connector-java	Extras	
Hive	Serviços	hadoop-httpfs	Extras	3.4.0.amzn.2-1.amzn2023
Hive	Serviços	hadoop-yarn-proxyserver	Extras	3.4.0.amzn.2-1.amzn2023
Hive	Serviços	hive-hbase	Extras	3.1.3.amzn.15-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Hive	Serviços	mariadb-connector-java	Extras	2.7.2-1
Hive	Serviços	s3-dist-cp	Extras	2.36.0-1.amzn2023
Phoenix	Serviços	s3-dist-cp	Extras	2.36.0-1.amzn2023
Spark	Serviços	spark-datanucleus	Extras	3.5.1.amzn.0-1.amzn2023
Tez	Serviços	hadoop-mapreduce	Extras	3.4.0.amzn.2-1.amzn2023
Tez	Serviços	hadoop-mapreduce-historyserver	Extras	3.4.0.amzn.2-1.amzn2023
Tez	Serviços	hadoop-yarn-proxyserver	Extras	3.4.0.amzn.2-1.amzn2023
Trino	Serviços	bigtop-utils	Extras	1.2.0-1.amzn2023
Trino	Serviços	hadoop-client	Extras	3.4.0.amzn.2-1.amzn2023
Trino	Serviços	hadoop-hdfs	Extras	3.4.0.amzn.2-1.amzn2023
Trino	Serviços	hadoop-hdfs-datanode	Extras	3.4.0.amzn.2-1.amzn2023
Trino	Serviços	hadoop-hdfs-journalnode	Extras	3.4.0.amzn.2-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Trino	Serviços	hadoop-hdfs-namenode	Extras	3.4.0.amzn.2-1.amzn2023
Trino	Serviços	hadoop-kms	Extras	3.4.0.amzn.2-1.amzn2023
Trino	Serviços	hadoop-mapreduce	Extras	3.4.0.amzn.2-1.amzn2023
Trino	Serviços	hadoop-mapreduce-historyserver	Extras	3.4.0.amzn.2-1.amzn2023
Trino	Serviços	hadoop-yarn	Extras	3.4.0.amzn.2-1.amzn2023
Trino	Serviços	hadoop-yarn-nodemanager	Extras	3.4.0.amzn.2-1.amzn2023
Trino	Serviços	hadoop-yarn-proxyserver	Extras	3.4.0.amzn.2-1.amzn2023
Trino	Serviços	hadoop-yarn-resourcemanager	Extras	3.4.0.amzn.2-1.amzn2023
Trino	Serviços	hadoop-yarn-timelineserver	Extras	3.4.0.amzn.2-1.amzn2023
Trino	Serviços	hive	Extras	3.1.3.amzn.12-1.amzn2023
Trino	Serviços	hive-hcatalog-server	Extras	3.1.3.amzn.12-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Trino	Serviços	mariadb-connector-java	Extras	2.7.2-1
Trino	Serviços	ranger-kms	Extras	2.0.0-1.amzn2023
Trino	Serviços	zookeeper	Extras	3.9.2.amzn.0-1.amzn2023
Trino	Serviços	zookeeper-server	Extras	3.9.2.amzn.0-1.amzn2023
Trino	Serviços	bigtop-groovy	Extras	2.5.4-1.amzn2023
Trino	Serviços	bigtop-jsvc	Extras	1.2.4-1.amzn2023
Trino	Serviços	hadoop	Extras	3.4.0.amzn.2-1.amzn2023
Trino	Serviços	hadoop-lzo	Extras	0.4.19-1.amzn2023
Trino	Serviços	hive-hcatalog	Extras	3.1.3.amzn.15-1.amzn2023
Trino	Serviços	hive-jdbc	Extras	3.1.3.amzn.12-1.amzn2023
Zookeeper	Serviços	emrfs	Extras	3.9.2.amzn.0-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Zookeeper	Serviços	hadoop-client	Extras	3.9.2.amzn.0-1.amzn2023
Zookeeper	Serviços	hadoop-hdfs	Extras	3.9.2.amzn.0-1.amzn2023
Zookeeper	Serviços	hadoop-hdfs-datanode	Extras	3.9.2.amzn.0-1.amzn2023
Zookeeper	Serviços	hadoop-hdfs-journalnode	Extras	3.9.2.amzn.0-1.amzn2023
Zookeeper	Serviços	hadoop-hdfs-namenode	Extras	3.9.2.amzn.0-1.amzn2023
Zookeeper	Serviços	hadoop-httpfs	Extras	3.9.2.amzn.0-1.amzn2023
Zookeeper	Serviços	hadoop-kms	Extras	3.9.2.amzn.0-1.amzn2023
Zookeeper	Serviços	hadoop-mapreduce	Extras	3.4.0.amzn.2-1.amzn2023
Zookeeper	Serviços	hadoop-mapreduce-historyserver	Extras	3.4.0.amzn.2-1.amzn2023
Zookeeper	Serviços	hadoop-yarn	Extras	3.4.0.amzn.2-1.amzn2023
Zookeeper	Serviços	hadoop-yarn-nodemanager	Extras	3.4.0.amzn.2-1.amzn2023
Zookeeper	Serviços	hadoop-yarn-proxyserver	Extras	3.4.0.amzn.2-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Zookeeper	Serviços	hadoop-yarn-resourcemanager	Extras	3.4.0.amzn.2-1.amzn2023
Zookeeper	Serviços	hadoop-yarn-timelineserver	Extras	3.4.0.amzn.2-1.amzn2023
Zookeeper	Serviços	ranger-kms	Extras	2.0.0-1.amzn2023

Notas de lançamento da versão 7.6.0

As notas de lançamento a seguir incluem informações sobre a versão 7.6.0 do Amazon EMR.

Novos recursos

- Atualizações da aplicação: Amazon EMR 7.6.0 application upgrades include HBase 2.6.1, Flink 1.20.0, Spark 3.5.3, and Trino 457.

Problemas conhecidos

- Quando o escalonamento gerenciado está ativado, o YARN ResourceManager (RM) enfrenta um problema crítico de impasse, fazendo com que ele pare de responder ao passar do estado DESCOMISSIONADO para DESCOMISSIONADO com operações simultâneas. Isso acontece principalmente em ambientes de grandes clusters. Quando isso ocorre, tanto as transições de estado do nó quanto as transições de estado do aplicativo falham no processamento, paralisando efetivamente a funcionalidade principal do. ResourceManager Como solução alternativa, reinicie o RM JVM usando o seguinte comando no nó primário:

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

- Quando um cluster do EMR é iniciado com agendadores YARN diferentes do `CapacityScheduler` ou `FairScheduler` (por exemplo, o Agendador FIFO), a `YARNMemoryAvailablePercentage` métrica não é enviada para. CloudWatch Esse problema afeta a redução de escala em clusters habilitados para escalabilidade

gerenciada. Também pode afetar o escalonamento gerenciado personalizado se depender da `YARNMemoryAvailablePercentage` métrica. Como solução alternativa, use o `CapacityScheduler` ou`FairScheduler`.

Alterações, melhorias e problemas resolvidos

- CloudWatch o agente é compatível com Runtime Role Clusters para EMR 7.6 e versões posteriores. Para obter mais informações, consulte o CloudWatch agente da Amazon.
- A partir da versão 7.6.0 do Amazon Elastic Map Reduce (EMR), o conector do sistema de arquivos S3A agora oferece suporte à criptografia do lado do cliente do Amazon S3.
- A partir do EMR 7.6.0, o TLS é suportado em endpoints. HMaster RegionServer Para obter mais informações, consulte a tabela a seguir sobre o suporte TLS com HBase: HBase.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250527.1	6.1.134-152.225.am zn2023	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Malásia), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			da Virgínia), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Malásia), Europa (Londres), Ásia-Pacífico (Melbourne), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (AWS GovCloud AWS GovCloud Oregon), México (Central) , Oeste dos EUA (Norte da Califórnia), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv), Ásia-Pacífico (Taipei), Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250512.(	6.1.134-152.225.am zn2023	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Hyderabad), Ásia-Pacífico (Jacarta)), Ásia-Pacífico (Malásia), Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Oriente Médio (Oriente Médio) (Oriente Médio (Oriente Médio)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Emirados Árabes Unidos), Canadá (Central) , Oeste do Canadá (Calgary), (Leste dos EUA), AWS GovCloud (AWS GovCloud Oeste dos EUA), China (Pequim), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250428.1	6.1.134-150.224.amzn2023	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Jacarta) Arta), Ásia-Pacífico (Malásia), Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Malásia) , Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (EAU), Canadá (Central) , Oeste do Canadá

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Calgary), (Leste dos EUA), AWS GovCloud (Oeste dos EUA), China (AWS GovCloud Pequim), China (Ningxia), Israel (Tel Aviv), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250414.(	6.1.132-147.221.am zn2023	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250331.(	6.1.131-143.221.am zn2023	18 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central), Ásia-Pacífico (Tóquio)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250303.(	6.1.129-138.220.am zn2023	27 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250218.1	6.1.128-136.201.amzn2023	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderabad), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250211.(	6.1.127-135.201.am zn2023	26 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250123.4	6.1.124-134.200.amzn2023	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderabad), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250115.(	6.1.119-129.201.am zn2023	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.5.2 240916	6.1.109-118.189.am zn2023	6 de setembro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

7.6.0 versões padrão do Java

O Amazon EMR versão 7.6 e superior vem com o Amazon Corretto 17 (construído no OpenJDK) por padrão para aplicativos que suportam o Corretto 17 (JDK 17), com exceção do Apache Livy.

A tabela a seguir mostra as versões padrão do Java para aplicativos no Amazon EMR 7.6.0. Se você quiser alterar a JVM padrão no cluster, siga as instruções em Configurar aplicações para usar uma Máquina Virtual Java específica para cada aplicação executada no cluster. Você só pode usar uma versão de runtime do Java para um cluster. O Amazon EMR não é compatível com a execução de diferentes nós ou aplicações em diferentes versões de runtime no mesmo cluster.

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
Delta	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Tez	17, 11, 8
Trino	21, 17
Zeppelin	17, 11, 8
Pig	17, 11, 8
Zookeeper	17, 11, 8

Versões de componentes 7.6.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado

`myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
adot-java-agent	1.31.0	Um Java Agent que coleta métricas de daemons de aplicações.
delta	3.2.1-amzn-1	O Delta Lake é um formato de tabela aberto para grandes conjuntos de dados analíticos
emr-amazon-cloudwatch-agent	1.300032.2-amzn-0	Um aplicativo que coleta métricas internas em nível de sistema e métricas de aplicativos personalizadas de instâncias da Amazon. EC2
emr-ddb	5.6.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.15.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.19.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.18.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.36.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	2.15.0	EMR S3Select Connector

Componente	Versão	Descrição
emr-wal-cli	1.3.1	Cli usada para lista/exclusão de emrwal.
emrfs	2.67,0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.20.0-amzn-0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.20.0-amzn-0	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
hadoop-client	3.4.0-aman-2	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	3.4.0-aman-2	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.4.0-aman-2	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.4.0-aman-2	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	3.4.0-aman-2	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.4.0-aman-2	O endpoint de HTTP para as operações do HDFS.

Componente	Versão	Descrição
hadoop-kms-server	3.4.0-aman-2	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	3.4.0-aman-2	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.4.0-aman-2	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	3.4.0-aman-2	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	3.4.0-aman-2	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	2.6.1-amzn-0	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	2.6.1-amzn-0	Serviço para atender a uma ou mais HBase regiões.
hbase-client	2.6.1-amzn-0	HBase cliente de linha de comando.

Componente	Versão	Descrição
hbase-rest-server	2.6.1-amzn-0	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.6.1-amzn-0	Serviço que fornece um endpoint Thrift para. HBase
hbase-operator-tools	2.6.1-amzn-0	Ferramenta de reparo para HBase clusters Apache.
hcatalog-client	3.1.3-amzn-15	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.3-amzn-15	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	3.1.3-amzn-15	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.3-amzn-15	O cliente da linha de comando do Hive.
hive-hbase	3.1.3-amzn-15	Cliente do Hive-hbase.
hive-metastore-server	3.1.3-amzn-15	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.

Componente	Versão	Descrição
hive-server2	3.1.3-amzn-15	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.15.0-aman-3	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.15.0-aman-3	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-trino	0.15.0-aman-3	Biblioteca de pacotes para execução do Presto com o Hudi.
hudi-spark	0.15.0-aman-3	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.11.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
iceberg	1.6.1-amzn-2	Apache Iceberg é um formato de tabela aberta para conjuntos de dados analíticos imensos
jupyterhub	1.5.0	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.8.0-incubating	Interface REST para interagir com o Apache Spark

Componente	Versão	Descrição
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.
nvidia-cuda	12.3.0	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.7.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	5.2.0	As bibliotecas do phoenix para servidor e cliente
phoenix-connectors	6.0.0	Conectores do Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0,287-amzn-1	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.

Componente	Versão	Descrição
presto-worker	0,287-amzn-1	O serviço que executa partes de uma consulta.
presto-client	0,287-amzn-1	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
trino-coordinator	457-amzn-0	Serviço que aceita consultas e gerencia a execução de consultas entre trino-workers.
trino-worker	457-amzn-0	O serviço que executa partes de uma consulta.
trino-client	457-amzn-0	Cliente de linha de comando do Trino instalado em mestres de espera de um cluster de alta disponibilidade (HA) em que o servidor do Trino não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.3.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	3.5.3-amzn-0	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	3.5.3-amzn-0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.5.3-amzn-0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.5.3-amzn-0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
spark-rapids	24.10.1-amzn-0	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs
tensorflow	2.16.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.10.2-amzn-13	O aplicativo e as bibliotecas do YARN para Tez.
tez-on-worker	0.10.2-amzn-13	A aplicação e as bibliotecas do YARN nós de processam ento.
zeppelin-server	0.11.1	O notebook baseado na web que permite um data analytics interativo.

Componente	Versão	Descrição
zookeeper-server	3.9.2-amzn-0	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.9.2-amzn-0	ZooKeeper cliente de linha de comando.

Classificações de configuração 7.6.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

classificações do emr-7.6.0

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Classificações	Descrição	Ações de reconfiguração
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.

Classificações	Descrição	Ações de reconfiguração
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Alterar a página inicial do KMS do Hadoop	Not available.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.

Classificações	Descrição	Ações de reconfiguração
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.	Not available.
iceberg-defaults	Alterar os valores no arquivo iceberg-defaults.conf do Iceberg.	Not available.
delta-defaults	Alterar os valores no arquivo delta-defaults.conf do Delta.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j2	Alterar as configurações de log4j2.properties do Livy.	Restarts Livy Server.

Classificações	Descrição	Ações de reconfiguração
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j2	Altere os valores no arquivo log4j2.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.

Classificações	Descrição	Ações de reconfiguração
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-lakeformation	Alterar os valores no arquivo lakeformation.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
trino-log	Alterar os valores no arquivo log.properties do Trino.	Restarts Trino-Server (for Trino)
trino-config	Alterar os valores no arquivo config.properties do Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Alterar os valores no arquivo password-authenticator.properties do Trino.	Restarts Trino-Server (for Trino)
trino-env	Alterar os valores no arquivo trino-env.sh do Trino.	Restarts Trino-Server (for Trino)
trino-node	Alterar os valores no arquivo node.properties do Trino.	Not available.
trino-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Trino.	Not available.
trino-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Trino.	Not available.
trino-connector-delta	Alterar os valores no arquivo delta.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Alterar os valores no arquivo hive.properties do Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Alterar os valores no arquivo exchange-manager.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Alterar os valores no arquivo iceberg.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Alterar os valores no arquivo hudi.properties do Trino.	Restarts Trino-Server (for Trino)

Classificações	Descrição	Ações de reconfiguração
trino-connector-jmx	Alterar os valores no arquivo jmx.properties do Trino.	Not available.
trino-connector-kafka	Alterar os valores no arquivo kafka.properties do Trino.	Not available.
trino-connector-localfile	Alterar os valores no arquivo localfile.properties do Trino.	Not available.
trino-connector-memory	Alterar os valores no arquivo memory.properties do Trino.	Not available.
trino-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Trino.	Not available.
trino-connector-mysql	Alterar os valores no arquivo mysql.properties do Trino.	Not available.
trino-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Trino.	Not available.
trino-connector-raptor	Alterar os valores no arquivo raptor.properties do Trino.	Not available.
trino-connector-redis	Alterar os valores no arquivo redis.properties do Trino.	Not available.
trino-connector-redshift	Alterar os valores no arquivo redshift.properties do Trino.	Not available.
trino-connector-tpch	Alterar os valores no arquivo tpch.properties do Trino.	Not available.
trino-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Trino.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Alterar os valores no arquivo kms-logback.xml do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j2	Alterar os valores no arquivo log4j2.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie and HiveServer2.

Classificações	Descrição	Ações de reconfiguração
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.
zookeeper-logback	Altere os valores no ZooKeeper arquivo logback.xml do.	Restarts Zookeeper server.
emr-metrics	Altere as configurações de métricas do EMR desse nó.	Restarts the CloudWatchAgent service.

Registro de alterações 7.6.0

Registro de alterações da versão 7.6.0 e das notas de lançamento

Data	Event	Descrição
2025-01-13	Publicação de documentos	Notas de lançamento do Amazon EMR 7.6.0 publicadas pela primeira vez
2025-01-10	Implantação concluída	Amazon EMR 7.6.0 totalmente implantado em todas as regiões suportadas
2025-12-26	Lançamento inicial	Amazon EMR 7.6.0 foi implantado pela primeira vez nas regiões comerciais iniciais

Amazon EMR versão 7.5.0

Ciclo de vida compatível com 7.5.0

A tabela a seguir descreve as datas de ciclo de vida suportadas para a versão 7.5.0 do Amazon EMR.

Fase de suporte	Data
Data da versão inicial	21 de novembro de 2024
Suporte padrão até	20 de novembro de 2026
Término do suporte	21 de novembro de 2026
Fim da vida útil	21 de novembro de 2027

Para obter uma lista de vulnerabilidades e exposições (CVEs) comuns críticas e altas que não afetam os clusters do EMR de acordo com o software e as configurações recomendados, consulte 7.5.0 vulnerabilidades e exposições comuns conhecidas para mecanismos principais.

Componentes da versão 7.5.0

O Amazon EMR 7.5.0 inclui os seguintes aplicativos:

Mecanismos centrais:

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

Extras:

Hue, Oozie JupyterHub, Porco, Presto, Zeppelin TensorFlow

Para obter mais detalhes sobre os pacotes incluídos, consulte Versões do aplicativo nas versões 7.x do Amazon EMR. Para obter mais informações sobre os mecanismos centrais e extras, consulte Amazon EMR Standard Support.

Versões do aplicativo 7.5.0

Essa versão inclui os seguintes aplicativos: AmazonCloudWatchAgent, Delta, Flink, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, Oozie, Phoenix, Pig, Presto, Spark, TensorFlow, TezTrino, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-7.5.0	emr-7.4.0	emr-7.3.0	emr-7.2.0
AWS SDK for Java	2.28.8, 1.12.772	2.28.8, 1.12.772	2.25.70, 1.12.747	2.23.18, 1.12.705

	emr-7.5.0	emr-7.4.0	emr-7.3.0	emr-7.2.0
Python	3,9, 3,11	3,9, 3,11	3,9, 3,11	3,9, 3,11
Scala	2.12.18	2.12.18	2.12.18	2.12.17
AmazonCloudWatchAgent	1.300032.2	1.300032.2	1.300032.2	1.300032.2
Delta	3.2.1-amzn-0	3.2.0-amzn-1	3.2.0-amzn-0	3.1.0
Flink	1.19.1	1.19.1	1.18.1	1.18.1
Ganglia	-	-	-	-
HBase	2.5.10	2.5.5	2.4.17	2.4.17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.4.0	3.4.0	3.3.6	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-amzn-2	0.15.0-amzn-1	0.15.0-amzn-0	0.14.1-amzn-1
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.6.1-amzn-1	1.6.1-amzn-0	1.5.2-amzn-0	1.5.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.8.0
MXNet	-	-	-	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1

	emr-7.5.0	emr-7.4.0	emr-7.3.0	emr-7.2.0
Phoenix	5.2.0	5.2.0	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,287	0,287	0,285	0,285
Spark	3.5.2	3.5.2	3.5.1	3.5.1
Sqoop	-	1.4.7	1.4.7	1.4.7
TensorFlow	2.16.1	2.16.1	2.16.1	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	446	446	442	436
Zeppelin	0.11.1	0.11.1	0.11.1	0.10.1
ZooKeeper	3.9.2	3.9.2	3.9.1	3.9.1

7.5.0 vulnerabilidades e exposições comuns

A tabela a seguir lista todos os CVEs que não afetam os clusters do EMR que são executados nas configurações recomendadas do Amazon EMR 7.5.0. O Amazon EMR depende do código upstream aberto para disponibilidade de correções e fornecerá a versão estável mais recente como parte dos componentes do mecanismo central em até 90 dias após a verificação das correções pelo Amazon EMR.

Essa tabela foi atualizada em 21 de novembro de 2024.

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-35728	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-35728

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2021-37533	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-3753
CVE-2018-14721	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-14 721
CVE-2020-26939	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-26939
CVE-2022-3171	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2024-47554	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-47 554
CVE-2021-38153	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-38 153
CVE-2020-36188	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-3618
CVE-2020-10673	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-10 673
CVE-2020-8908	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-8908
CVE-2022-3509	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-41966	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-4 1966
CVE-2023-3635	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-3635
CVE-2020-25649	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-25 649

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2017-17485	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2017-17485
CVE-2024-25710	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-25710
CVE-2023-25194	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-25194
GHSA-58QW-P7QM-5RVh	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/GHSA-58qw-p7qm-5rvh
CVE-2023-51074	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-510 74
CVE-2023-20883	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-2083
CVE-2022-42004	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2004
CVE-2024-47561	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-47 56
CVE-2023-2976	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2022-45685	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-45 685
CVE-2023-33202	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-33202
CVE-2019-14892	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-14 892
CVE-2023-33201	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-33201

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2023-1436	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2018-10237	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-10 237
CVE-2015-5237	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2015-5237
CVE-2022-42003	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2003
CVE-2024-36114	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-3614
CVE-2024-7254	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-7254
CVE-2020-14061	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14061
CVE-2024-23944	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-2394
CVE-2020-10969	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10 969
CVE-2023-52428	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-52428
CVE-2024-29857	BAIXO	https://nvd.nist. gov/vuln/detail/CVE-2024-29 857
CVE-2024-30171	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-30 171
CVE-2022-41854	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-4 1854

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2022-3510	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3510
CVE-2018-7489	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-7 489
CVE-2020-35491	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-35491
CVE-2020-11111	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-1111
CVE-2020-15250	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-15250
CVE-2020-9548	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-9548
CVE-2024-51504	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-51504
CVE-2023-6378	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-6378
CVE-2023-34453	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3445 3
CVE-2024-29025	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-29 025
CVE-2024-36124	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-36124
CVE-2013-4002	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2013-4002
CVE-2019-14379	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14 379

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-36186	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-36186
CVE-2022-45688	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-4568
CVE-2020-17521	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-17 521
CVE-2024-8184	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-8184
CVE-2020-14195	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-14 195
CVE-2023-40167	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-40167
CVE-2024-6763	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-6763
CVE-2022-25647	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2022-25 647
CVE-2018-14719	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2018-14 719
CVE-2024-29131	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-29 131
CVE-2020-9546	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-9546
CVE-2018-1000180	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-1000180
CVE-2023-39410	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-39410

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2021-22570	BAIXO	https://nvd.nist. gov/vuln/detail/CVE-2021-22570
CVE-2022-46751	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2022-46751
CVE-2021-36373	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2021-36373
CVE-2023-26048	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-26048
CVE-2024-30172	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-30172
CVE-2022-36364	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3636 4
CVE-2022-46337	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-46337
CVE-2019-16869	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-16 869
CVE-2022-38751	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-38751
CVE-2022-1415	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-1415
CVE-2021-22569	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-22569
CVE-2024-35255	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-3525
CVE-2021-28170	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-28 170

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2022-40149	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2022-40149
CVE-2023-50868	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-50868
CVE-2023-34454	high	https://nvd.nist. gov/vuln/detail/CVE-2023-3445
CVE-2022-38750	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2022-38750
CVE-2020-14060	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14060
CVE-2023-42503	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-42503
CVE-2024-25638	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-25 638
CVE-2024-29133	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-2913
CVE-2022-37865	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-37865
CVE-2018-19361	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-1936 1
CVE-2018-19362	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2018-1936 2
CVE-2020-10650	moderar	https://nvd.nist. gov/vuln/detail/CVE-2020-10650
CVE-2020-35490	moderar	https://nvd.nist. gov/vuln/detail/CVE-2020-35490

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2024-23454	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-23454
CVE-2023-43642	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2020-24750	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-24750
CVE-2020-11112	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-1112
CVE-2019-14439	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14 439
CVE-2023-34455	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-3445
CVE-2024-26308	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-26308
CVE-2019-20445	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-20 445
CVE-2023-0833	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-0833
CVE-2020-36185	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36185
CVE-2023-6481	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-6481
CVE-2022-38752	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2022-38752
CVE-2020-1945	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-1945

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-13936	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-13936
CVE-2023-36478	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-36478
CVE-2020-10968	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-10 968
CVE-2020-28052	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-28052
CVE-2021-43859	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-438 59
CVE-2021-27568	BAIXO	https://nvd.nist. gov/vuln/detail/CVE-2021-27568
CVE-2020-11113	BAIXO	https://nvd.nist. gov/vuln/detail/CVE-2020-11-13
CVE-2022-40150	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40150
CVE-2023-5072	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-5072
CVE-2018-11307	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2018-11 307
CVE-2024-21634	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-21634
CVE-2020-36518	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36518
CVE-2020-36180	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2020-36180

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2021-34538	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-34 538
CVE-2021-29425	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-29425
CVE-2020-8840	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-8840
CVE-2020-9547	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-9547
CVE-2020-36179	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36179
CVE-2019-16943	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2019-16 943
CVE-2016-5018	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2016-5018
CVE-2024-34447	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-3447
CVE-2024-39676	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-39676
CVE-2021-21290	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-21290
CVE-2019-12086	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-12 2086
CVE-2018-14718	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14 718
CVE-2020-36181	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36181

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2022-39135	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2022-39135
CVE-2019-16335	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2019-16 335
CVE-2020-36184	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36184
CVE-2022-45693	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-45693
CVE-2020-14062	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-14062
CVE-2020-15522	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-1552
CVE-2019-14540	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-14 540
CVE-2017-15095	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2017-15095
CVE-2018-5968	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-5968
CVE-2019-14893	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-14 893
CVE-2019-12814	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-12814
CVE-2020-10672	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-10 672
CVE-2020-24616	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-24 616

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-13949	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2023-50387	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-50387
CVE-2018-12022	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE- 2018-12022
CVE-2020-36189	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-36189
CVE-2021-37137	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-37137
CVE-2019-12384	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-12 384
CVE-2023-26049	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2023-26049
CVE-2022-1471	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-1471
CVE-2023-1370	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2019-16942	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-16 942
CVE-2022-40152	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40152
CVE-2023-20863	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-20863
CVE-2018-11798	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-17 98

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2019-20330	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-20330
CVE-2019-17531	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-17 531
CVE-2020-11620	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-11620
CVE-2021-31684	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-3 684
CVE-2020-11619	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-11619
CVE-2022-40151	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40151
CVE-2019-17267	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-17 267
CVE-2020-36182	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36182
CVE-2016-6796	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2016-6796
CVE-2019-0205	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-0205
CVE-2018-10054	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-10 054
CVE-2018-14720	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2018-14 720
CVE-2022-24329	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2022-24 329

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-36183	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-36183
CVE-2023-44981	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2022-37866	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-3786
CVE-2015-1832	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2015-1832
CVE-2023-38647	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-38647
CVE-2018-1313	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2018-13 13
CVE-2018-12023	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-12 023
CVE-2021-20190	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-2019 0
CVE-2024-38808	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-38808
CVE-2019-20444	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-20 44
CVE-2012-0881	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2012-08 81
CVE-2020-36187	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36187
CVE-2020-13956	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-13956

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2022-38749	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-38749
CVE-2017-7525	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2017-7525
CVE-2018-19360	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2018-19360
CVE-2021-37136	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2021-37136
CVE-2022-42889	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-4289
CVE-2021-36374	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-36374
CVE-2020-13955	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-1395
CVE-2022-25857	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-25 857

Detalhes dos componentes da versão 7.5.0

Consulte a tabela a seguir para obter mais informações sobre os pacotes extras no Amazon EMR 7.5.0.

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
AmazonCloudWatchAgent	Serviços	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.5-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
AmazonCloudWatchAgent	Serviços	ranger-kms	Extras	2.0.0-1.amzn2023
AmazonCloudWatchAgent	Serviços	s3-dist-cp	Extras	2.33.0-1.amzn2023
AmazonCloudWatchAgent	Serviços	zookeeper	Extras	3.9.1.amzn.1-1.amzn2023
AmazonCloudWatchAgent	Serviços	zookeeper-server	Extras	3.9.1.amzn.1-1.amzn2023
Flink	Serviços	hadoop-client	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-hdfs	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-hdfs-datanode	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-hdfs-namenode	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-httpfs	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-kms	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-mapreduce	Extras	3.3.6.amzn.5-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Flink	Serviços	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-yarn	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-yarn-nodemanager	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-yarn-resourcemanager	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-yarn-timelineserver	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hudi	Extras	0.15.0.amzn.0-1.amzn2023
Flink	Serviços	ranger-kms	Extras	2.0.0-1.amzn2023
Flink	Serviços	hadoop	Extras	3.3.6.amzn.5-1.amzn2023
HBase	Serviços	s3-dist-cp	Extras	2.33.0-1.amzn2023
HCatalog	Serviços	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
HCatalog	Serviços	mariadb-connector-java	Extras	
Hive	Serviços	hadoop-httpfs	Extras	3.3.6.amzn.5-1.amzn2023
Hive	Serviços	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023
Hive	Serviços	hive-hbase	Extras	3.1.3.amzn.12-1.amzn2023
Hive	Serviços	mariadb-connector-java	Extras	2.7.2-1
Hive	Serviços	s3-dist-cp	Extras	2.33.0-1.amzn2023
Phoenix	Serviços	s3-dist-cp	Extras	2.33.0-1.amzn2023
Spark	Serviços	spark-datanucleus	Extras	3.5.1.amzn.0-1.amzn2023
Tez	Serviços	hadoop-mapreduce	Extras	3.3.6.amzn.5-1.amzn2023
Tez	Serviços	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.5-1.amzn2023
Tez	Serviços	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Trino	Serviços	bigtop-utils	Extras	1.2.0-1.amzn2023
Trino	Serviços	hadoop-client	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-hdfs	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-hdfs-datanode	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-hdfs-namenode	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-kms	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-mapreduce	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-yarn	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-yarn-nodemanager	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Trino	Serviços	hadoop-yarn-resourcemanager	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-yarn-timelineserver	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hive	Extras	3.1.3.amzn.12-1.amzn2023
Trino	Serviços	hive-hcatalog-server	Extras	3.1.3.amzn.12-1.amzn2023
Trino	Serviços	mariadb-connector-java	Extras	2.7.2-1
Trino	Serviços	ranger-kms	Extras	2.0.0-1.amzn2023
Trino	Serviços	zookeeper	Extras	3.9.1.amzn.1-1.amzn2023
Trino	Serviços	zookeeper-server	Extras	3.9.1.amzn.1-1.amzn2023
Trino	Serviços	bigtop-groovy	Extras	2.5.4-1.amzn2023
Trino	Serviços	bigtop-jsvc	Extras	1.2.4-1.amzn2023
Trino	Serviços	hadoop	Extras	3.3.6.amzn.5-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Trino	Serviços	hadoop-lzo	Extras	0.4.19-1.amzn2023
Trino	Serviços	hive-hcatalog	Extras	3.1.3.amzn.12-1.amzn2023
Trino	Serviços	hive-jdbc	Extras	3.1.3.amzn.12-1.amzn2023
Zookeeper	Serviços	emrfs	Extras	2.64.0-1.amzn2023
Zookeeper	Serviços	hadoop-client	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-hdfs	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-hdfs-datanode	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-hdfs-namenode	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-httpfs	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-kms	Extras	3.3.6.amzn.5-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Zookeeper	Serviços	hadoop-mapreduce	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-yarn	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-yarn-nodemanager	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-yarn-resourcemanager	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-yarn-timelineserver	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	ranger-kms	Extras	2.0.0-1.amzn2023

Notas de lançamento da versão 7.5.0

As notas de lançamento a seguir incluem informações sobre a versão 7.5.0 do Amazon EMR.

Novos recursos

- Atualizações da aplicação: Amazon EMR 7.5.0 application upgrades include Delta 3.2.1-amzn-0, HBase 2.5.10, Hudi 0.15.0-amzn-2, and Iceberg 1.6.1-amzn-1.
- Os registros de gravação antecipada (WAL) do EMR oferecem suporte à replicação entre clusters. Para obter mais informações, consulte Replicação entre clusters do EMR WAL.

Problemas conhecidos

- O Apache Sqoop foi removido do Amazon EMR versão 7.5 e versões posteriores.
- Quando o escalonamento gerenciado está ativado, o YARN ResourceManager (RM) enfrenta um problema crítico de impasse, fazendo com que ele pare de responder ao passar do estado DESCOMISSIONADO para DESCOMISSIONADO com operações simultâneas. Isso acontece principalmente em ambientes de grandes clusters. Quando isso ocorre, tanto as transições de estado do nó quanto as transições de estado do aplicativo falham no processamento, paralisando efetivamente a funcionalidade principal do. ResourceManager Como solução alternativa, reinicie o RM JVM usando o seguinte comando no nó primário:

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

- Quando um cluster do EMR é iniciado com agendadores YARN diferentes do `CapacityScheduler` ou `FairScheduler` (por exemplo, o Agendador FIFO), a `YARNMemoryAvailablePercentage` métrica não é enviada para. CloudWatch Esse problema afeta a redução de escala em clusters habilitados para escalabilidade gerenciada. Também pode afetar o escalonamento gerenciado personalizado se depender da `YARNMemoryAvailablePercentage` métrica. Como solução alternativa, use o `CapacityScheduler` ou`FairScheduler`.

Alterações, melhorias e problemas resolvidos

- O Amazon EMR usa diferentes versões da TensorFlow biblioteca, dependendo dos tipos de instância que você escolher para o seu cluster. O Amazon EMR também oferece suporte TensorFlow para clusters com tipos de instância aarch64 (instâncias Graviton) do EMR-7.5.0 e superior.

  A tabela a seguir lista as compilações por tipo de instância para o EMR-7.5.0:

EC2 tipos de instância	TensorFlow construir		
M5 e C5	Tensorflow 2.16.1 com otimização do Intel MKL		

EC2 tipos de instância	TensorFlow construir		
P2, P4D, P5, G4DN, G5, G6 e GR6	Tensorflow 2.16.1 com CUDA 12.3, cuDNN 8.9.7.29		
P3, P3DN, G3 e G3S	O Tensorflow 2.16.1 com CUDA 12.3, cuDNN 8.9.7.29, NCCL 2.20.3-1 A Nvidia NCCL está disponível somente em instância s P3. Contrato de licença de usuário final (EULA) — Ao usar componentes da Nvidia no Amazon EMR, você concorda com os termos e condições descritos no EULA do produto.		
Todos os outros, incluindo Graviton (EMR-7.5.0+)	Tensorflow 2.16.1		

- Um cluster do Amazon EMR para AWS Lake Formation com FGAC (controle de acesso refinado) habilitado agora pode suportar até 256 etapas simultâneas com a seguinte configuração de hdfs-site.xml:

```
[
  {
    "Classification": "hdfs-site",
    "Properties": {
      "dfs.datanode.handler.count": "40",
      "dfs.datanode.max.transfer.threads": "16384",
      "dfs.namenode.handler.count": "64",
```

```
      "dfs.replication": "1"
    }
  },
  {
    "Classification": "hdfs-env",
    "Configurations": [
      {
        "Classification": "export",
        "Properties": {
          "HADOOP_DATANODE_HEAPSIZE": "5120"
        }
      }
    ],
    "Properties": {}
  }
]
```

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250527.1	6.1.134-152.225.amzn2023	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein), China (Pequim), Ásia-Pacífico (Malásia), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Leste dos EUA (Norte da Virgínia), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Malásia), Europa (Londres), Ásia-Pacífico (Melbourne), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (AWS GovCloud AWS GovCloud Oregon), México (Central), Oeste dos EUA (Norte da Califórnia), Ásia-Pacífico

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv), Ásia-Pacífico (Taipei), Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250512.(	6.1.134-152.225.am zn2023	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Hyderabad), Ásia-Pacífico (Jacarta)), Ásia-Pacífico (Malásia), Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Oriente Médio (Oriente Médio) (Oriente Médio (Oriente Médio)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Emirados Árabes Unidos), Canadá (Central) , Oeste do Canadá (Calgary), (Leste dos EUA), AWS GovCloud (AWS GovCloud Oeste dos EUA), China (Pequim), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250428.1	6.1.134-150.224.amzn2023	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Jacarta) Arta), Ásia-Pacífico (Malásia), Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Malásia) , Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (EAU), Canadá (Central) , Oeste do Canadá

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Calgary), (Leste dos EUA), AWS GovCloud (Oeste dos EUA), China (AWS GovCloud Pequim), China (Ningxia), Israel (Tel Aviv), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250414.(	6.1.132-147.221.am zn2023	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250331.(	6.1.131-143.221.am zn2023	18 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central), Ásia-Pacífico (Tóquio)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250303.(	6.1.129-138.220.am zn2023	27 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250218.1	6.1.128-136.201.amzn2023	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderabad), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250211.(	6.1.127-135.201.am zn2023	26 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250123.4	6.1.124-134.200.amzn2023	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderabad), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250115.(	6.1.119-129.201.am zn2023	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.5.2 240916	6.1.109-118.189.am zn2023	6 de setembro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

Versões Java padrão 7.5.0

O Amazon EMR versão 7.5 e superior vem com o Amazon Corretto 17 (construído no OpenJDK) por padrão para aplicativos que suportam o Corretto 17 (JDK 17), com exceção do Apache Livy.

A tabela a seguir mostra as versões padrão do Java para aplicativos no Amazon EMR 7.5.0. Se você quiser alterar a JVM padrão no cluster, siga as instruções em Configurar aplicações para usar uma Máquina Virtual Java específica para cada aplicação executada no cluster. Você só pode usar uma versão de runtime do Java para um cluster. O Amazon EMR não é compatível com a execução de diferentes nós ou aplicações em diferentes versões de runtime no mesmo cluster.

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
Delta	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Tez	17, 11, 8
Trino	21, 17
Zeppelin	17, 11, 8
Pig	8
Zookeeper	17, 11, 8

Versões de componentes 7.5.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado

`myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
adot-java-agent	1.31.0	Um Java Agent que coleta métricas de daemons de aplicações.
delta	3.2.1-amzn-0	O Delta Lake é um formato de tabela aberto para grandes conjuntos de dados analíticos
emr-amazon-cloudwatch-agent	1.300032.2-amzn-0	Um aplicativo que coleta métricas internas em nível de sistema e métricas de aplicativos personalizadas de instâncias da Amazon. EC2
emr-ddb	5.6.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.14.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.18.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.16.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.35.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	2.14.0	EMR S3Select Connector

Componente	Versão	Descrição
emr-wal-cli	1.3.1	Cli usada para lista/exclusão de emrwal.
emrfs	2.66.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.19.1-amzn-1	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.19.1-amzn-1	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
hadoop-client	3.4.0-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	3.4.0-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.4.0-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.4.0-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	3.4.0-amzn-1	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.4.0-amzn-1	O endpoint de HTTP para as operações do HDFS.

Componente	Versão	Descrição
hadoop-kms-server	3.4.0-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	3.4.0-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.4.0-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	3.4.0-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	3.4.0-amzn-1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	2.5.10-amzn-0	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	2.5.10-amzn-0	Serviço para atender a uma ou mais HBase regiões.
hbase-client	2.5.10-amzn-0	HBase cliente de linha de comando.

Componente	Versão	Descrição
hbase-rest-server	2.5.10-amzn-0	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.5.10-amzn-0	Serviço que fornece um endpoint Thrift para. HBase
hbase-operator-tools	2.5.10-amzn-0	Ferramenta de reparo para HBase clusters Apache.
hcatalog-client	3.1.3-aman-14	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.3-aman-14	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	3.1.3-aman-14	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.3-aman-14	O cliente da linha de comando do Hive.
hive-hbase	3.1.3-aman-14	Cliente do Hive-hbase.
hive-metastore-server	3.1.3-aman-14	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.

Componente	Versão	Descrição
hive-server2	3.1.3-aman-14	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.15.0-amzn-2	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.15.0-amzn-2	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-trino	0.15.0-amzn-2	Biblioteca de pacotes para execução do Presto com o Hudi.
hudi-spark	0.15.0-amzn-2	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.11.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
iceberg	1.6.1-amzn-1	Apache Iceberg é um formato de tabela aberta para conjuntos de dados analíticos imensos
jupyterhub	1.5.0	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.8.0-incubating	Interface REST para interagir com o Apache Spark

Componente	Versão	Descrição
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.
nvidia-cuda	12.3.0	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.7.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	5.2.0	As bibliotecas do phoenix para servidor e cliente
phoenix-connectors	6.0.0	Conectores do Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0,287-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.

Componente	Versão	Descrição
presto-worker	0,287-amzn-0	O serviço que executa partes de uma consulta.
presto-client	0,287-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
trino-coordinator	446-aman-0	Serviço que aceita consultas e gerencia a execução de consultas entre trino-workers.
trino-worker	446-aman-0	O serviço que executa partes de uma consulta.
trino-client	446-aman-0	Cliente de linha de comando do Trino instalado em mestres de espera de um cluster de alta disponibilidade (HA) em que o servidor do Trino não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.3.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	3.5.2-amzn-1	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	3.5.2-amzn-1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.5.2-amzn-1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.5.2-amzn-1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
spark-rapids	24.08.1-amzn-1	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs
tensorflow	2.16.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.10.2-amzn-12	O aplicativo e as bibliotecas do YARN para Tez.
tez-on-worker	0.10.2-amzn-12	A aplicação e as bibliotecas do YARN nós de processam ento.
zeppelin-server	0.11.1	O notebook baseado na web que permite um data analytics interativo.

Componente	Versão	Descrição
zookeeper-server	3.9.2-amzn-0	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.9.2-amzn-0	ZooKeeper cliente de linha de comando.

Classificações de configuração 7.5.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

classificações do emr-7.5.0

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Classificações	Descrição	Ações de reconfiguração
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.

Classificações	Descrição	Ações de reconfiguração
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Alterar a página inicial do KMS do Hadoop	Not available.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.

Classificações	Descrição	Ações de reconfiguração
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.	Not available.
iceberg-defaults	Alterar os valores no arquivo iceberg-defaults.conf do Iceberg.	Not available.
delta-defaults	Alterar os valores no arquivo delta-defaults.conf do Delta.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j2	Alterar as configurações de log4j2.properties do Livy.	Restarts Livy Server.

Classificações	Descrição	Ações de reconfiguração
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j2	Altere os valores no arquivo log4j2.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.

Classificações	Descrição	Ações de reconfiguração
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-lakeformation	Alterar os valores no arquivo lakeformation.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
trino-log	Alterar os valores no arquivo log.properties do Trino.	Restarts Trino-Server (for Trino)
trino-config	Alterar os valores no arquivo config.properties do Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Alterar os valores no arquivo password-authenticator.properties do Trino.	Restarts Trino-Server (for Trino)
trino-env	Alterar os valores no arquivo trino-env.sh do Trino.	Restarts Trino-Server (for Trino)
trino-node	Alterar os valores no arquivo node.properties do Trino.	Not available.
trino-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Trino.	Not available.
trino-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Trino.	Not available.
trino-connector-delta	Alterar os valores no arquivo delta.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Alterar os valores no arquivo hive.properties do Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Alterar os valores no arquivo exchange-manager.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Alterar os valores no arquivo iceberg.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Alterar os valores no arquivo hudi.properties do Trino.	Restarts Trino-Server (for Trino)

Classificações	Descrição	Ações de reconfiguração
trino-connector-jmx	Alterar os valores no arquivo jmx.properties do Trino.	Not available.
trino-connector-kafka	Alterar os valores no arquivo kafka.properties do Trino.	Not available.
trino-connector-localfile	Alterar os valores no arquivo localfile.properties do Trino.	Not available.
trino-connector-memory	Alterar os valores no arquivo memory.properties do Trino.	Not available.
trino-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Trino.	Not available.
trino-connector-mysql	Alterar os valores no arquivo mysql.properties do Trino.	Not available.
trino-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Trino.	Not available.
trino-connector-raptor	Alterar os valores no arquivo raptor.properties do Trino.	Not available.
trino-connector-redis	Alterar os valores no arquivo redis.properties do Trino.	Not available.
trino-connector-redshift	Alterar os valores no arquivo redshift.properties do Trino.	Not available.
trino-connector-tpch	Alterar os valores no arquivo tpch.properties do Trino.	Not available.
trino-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Trino.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Alterar os valores no arquivo kms-logback.xml do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j2	Alterar os valores no arquivo log4j2.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie and HiveServer2.

Classificações	Descrição	Ações de reconfiguração
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.
zookeeper-logback	Altere os valores no ZooKeeper arquivo logback.xml do.	Restarts Zookeeper server.
emr-metrics	Altere as configurações de métricas do EMR desse nó.	Restarts the CloudWatchAgent service.

Registro de alterações 7.5.0

Registro de alterações da versão 7.5.0 e das notas de lançamento

Data	Event	Descrição
2024-11-21	Publicação de documentos	Notas de lançamento do Amazon EMR 7.5.0 publicadas pela primeira vez
2024-11-21	Implantação concluída	Amazon EMR 7.5.0 totalmente implantado em todas as regiões suportadas
2024-11-13	Lançamento inicial	Amazon EMR 7.5.0 implantado pela primeira vez nas regiões comerciais iniciais

Amazon EMR versão 7.4.0

Ciclo de vida compatível com 7.4.0

A tabela a seguir descreve as datas de ciclo de vida suportadas para a versão 7.4.0 do Amazon EMR.

Fase de suporte	Data
Data da versão inicial	13 de novembro de 2024
Suporte padrão até	12 de novembro de 2026
Término do suporte	13 de novembro de 2026
Fim da vida útil	13 de novembro de 2027

Para obter uma lista de vulnerabilidades e exposições (CVEs) comuns críticas e altas que não afetam os clusters do EMR de acordo com o software e as configurações recomendados, consulte 7.4.0 vulnerabilidades e exposições comuns conhecidas para mecanismos principais.

Componentes da versão 7.4.0

O Amazon EMR 7.4.0 inclui os seguintes aplicativos:

Mecanismos centrais:

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

Extras:

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

Para obter mais detalhes sobre os pacotes incluídos, consulte Versões do aplicativo nas versões 7.x do Amazon EMR. Para obter mais informações sobre os mecanismos centrais e extras, consulte Amazon EMR Standard Support.

Versões do aplicativo 7.4.0

Essa versão inclui os seguintes aplicativos: AmazonCloudWatchAgent, Delta, Flink, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, Oozie, Phoenix, Pig, Presto, Spark, SqoopTensorFlow, Tez, Trino, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-7.4.0	emr-7.3.0	emr-7.2.0	emr-7.1.0
AWS SDK for Java	2.28.8, 1.12.772	2.25.70, 1.12.747	2.23.18, 1.12.705	2.23.18, 1.12.656

	emr-7.4.0	emr-7.3.0	emr-7.2.0	emr-7.1.0
Python	3,9, 3,11	3,9, 3,11	3,9, 3,11	3,9, 3,11
Scala	2.12.18	2.12.18	2.12.17	2.12.17
AmazonCloudWatchAgent	1.300032.2	1.300032.2	1.300032.2	1.300032.2
Delta	3.2.0-amzn-1	3.2.0-amzn-0	3.1.0	3.0.0
Flink	1.19.1	1.18.1	1.18.1	1.18.1
Ganglia	-	-	-	-
HBase	2.5.5	2.4.17	2.4.17	2.4.17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.4.0	3.3.6	3.3.6	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-amzn-1	0.15.0-amzn-0	0.14.1-amzn-1	0.14.1-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.6.1-amzn-0	1.5.2-amzn-0	1.5.0-amzn-0	1.4.3-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.8.0
MXNet	-	-	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1

	emr-7.4.0	emr-7.3.0	emr-7.2.0	emr-7.1.0
Phoenix	5.2.0	5.1.3	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,287	0,285	0,285	0,284
Spark	3.5.2	3.5.1	3.5.1	3.5.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.16.1	2.16.1	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	446	442	436	435
Zeppelin	0.11.1	0.11.1	0.10.1	0.10.1
ZooKeeper	3.9.2	3.9.1	3.9.1	3.9.1

7.4.0 vulnerabilidades e exposições comuns

A tabela a seguir lista todos os CVEs que não afetam os clusters do EMR que são executados nas configurações recomendadas do Amazon EMR 7.4.0. O Amazon EMR depende do código upstream aberto para disponibilidade de correções e fornecerá a versão estável mais recente como parte dos componentes do mecanismo central em até 90 dias após a verificação das correções pelo Amazon EMR.

Essa tabela foi atualizada em 15 de novembro de 2024.

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2021-37533	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-3753

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2018-14721	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-14 721
CVE-2020-26939	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-26939
CVE-2022-3171	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2024-47554	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-47 554
CVE-2021-38153	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-38 153
CVE-2022-41966	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-4 1966
CVE-2023-3635	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-3635
CVE-2020-25649	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-25 649
CVE-2017-17485	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2017-17485
CVE-2024-25710	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-25710
CVE-2023-25194	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-25194
GHSA-58QW-P7QM-5RVh	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/GHSA-58qw-p7qm-5rvh
CVE-2023-51074	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-510 74

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2023-20883	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-2083
CVE-2022-42004	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2004
CVE-2024-47561	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-47 56
CVE-2023-2976	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2022-45685	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-45 685
CVE-2023-33202	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-33202
CVE-2019-14892	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-14 892
CVE-2023-33201	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-33201
CVE-2023-1436	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2018-10237	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-10 237
CVE-2015-5237	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2015-5237
CVE-2022-42003	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2003
CVE-2024-36114	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-3614

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2024-7254	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-7254
CVE-2020-14061	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14061
CVE-2024-23944	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-2394
CVE-2020-10969	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10 969
CVE-2023-52428	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-52428
CVE-2024-29857	BAIXO	https://nvd.nist. gov/vuln/detail/CVE-2024-29 857
CVE-2024-30171	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-30 171
CVE-2022-41854	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-4 1854
CVE-2022-3510	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3510
CVE-2018-7489	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-7 489
CVE-2020-35491	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-35491
CVE-2020-11111	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-1111
CVE-2020-15250	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-15250

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-9548	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-9548
CVE-2024-51504	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-51504
CVE-2023-6378	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-6378
CVE-2023-34453	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3445 3
CVE-2024-29025	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-29 025
CVE-2024-36124	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-36124
CVE-2013-4002	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2013-4002
CVE-2019-14379	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14 379
CVE-2020-36186	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-36186
CVE-2022-45688	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-4568
CVE-2020-17521	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-17 521
CVE-2024-8184	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-8184
CVE-2020-14195	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-14 195

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2023-40167	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-40167
CVE-2024-6763	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-6763
CVE-2022-25647	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2022-25 647
CVE-2018-14719	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2018-14 719
CVE-2024-29131	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-29 131
CVE-2020-9546	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-9546
CVE-2018-1000180	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-1000180
CVE-2023-39410	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2021-22570	BAIXO	https://nvd.nist. gov/vuln/detail/CVE-2021-22570
CVE-2022-46751	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2022-46751
CVE-2021-36373	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2021-36373
CVE-2023-26048	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-26048
CVE-2024-30172	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-30172

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2022-36364	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3636 4
CVE-2022-46337	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-4637
CVE-2019-16869	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-16 869
CVE-2022-38751	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-38751
CVE-2022-1415	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-1415
CVE-2021-22569	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-22569
CVE-2024-35255	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-3525
CVE-2021-28170	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-28 170
CVE-2022-40149	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2022-40149
CVE-2023-50868	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-50868
CVE-2022-38750	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2022-38750
CVE-2020-14060	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14060
CVE-2023-42503	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-42503

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2024-25638	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-25 638
CVE-2024-29133	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-2913
CVE-2022-37865	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-37865
CVE-2018-19361	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-1936 1
CVE-2018-19362	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2018-1936 2
CVE-2020-10650	moderar	https://nvd.nist. gov/vuln/detail/CVE-2020-10650
CVE-2020-35490	moderar	https://nvd.nist. gov/vuln/detail/CVE-2020-35490
CVE-2024-23454	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-23454
CVE-2023-43642	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2020-24750	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-24750
CVE-2020-11112	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-1112
CVE-2019-14439	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14 439
CVE-2023-34455	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-3445

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2024-26308	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-26308
CVE-2019-20445	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-20 445
CVE-2023-0833	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-0833
CVE-2020-36185	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36185
CVE-2023-6481	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-6481
CVE-2012-5783	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2012-5783
CVE-2022-38752	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2022-38752
CVE-2020-1945	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-1945
CVE-2020-13936	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-13936
CVE-2023-36478	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-36478
CVE-2020-10968	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-10 968
CVE-2020-28052	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-28052
CVE-2021-43859	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-438 59

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2021-27568	BAIXO	https://nvd.nist. gov/vuln/detail/CVE-2021-27568
CVE-2020-11113	BAIXO	https://nvd.nist. gov/vuln/detail/CVE-2020-11-13
CVE-2022-40150	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40150
CVE-2023-5072	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-5072
CVE-2018-11307	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-11 307
CVE-2019-16943	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16 943
CVE-2022-39135	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-39135
CVE-2019-16335	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16 335
CVE-2017-7525	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2017-7525
CVE-2018-19360	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-19360
CVE-2020-35728	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-35728
CVE-2023-34454	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3445
CVE-2024-21634	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-21634

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-36518	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36518
CVE-2012-6153	BAIXO	https://nvd.nist. gov/vuln/detail/CVE-2012-6153
CVE-2020-36180	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2020-36180
CVE-2021-34538	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-34 538
CVE-2021-29425	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-29425
CVE-2020-8840	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-8840
CVE-2020-9547	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-9547
CVE-2020-36179	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36179
CVE-2016-5018	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2016-5018
CVE-2024-34447	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-3447
CVE-2024-39676	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-39676
CVE-2021-21290	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-21290
CVE-2019-12086	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-12 2086

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2018-14718	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14 718
CVE-2020-36181	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36181
CVE-2020-36184	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36184
CVE-2022-45693	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-45693
CVE-2020-14062	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-14062
CVE-2020-15522	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-1552
CVE-2019-14540	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-14 540
CVE-2017-15095	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2017-15095
CVE-2018-5968	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-5968
CVE-2019-14893	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-14 893
CVE-2019-12814	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-12814
CVE-2020-10672	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-10 672
CVE-2020-24616	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-24 616

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-13949	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2023-50387	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-50387
CVE-2018-12022	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE- 2018-12022
CVE-2020-36189	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-36189
CVE-2021-37137	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-37137
CVE-2019-12384	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-12 384
CVE-2023-26049	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2023-26049
CVE-2022-1471	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-1471
CVE-2023-1370	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2019-16942	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-16 942
CVE-2022-40152	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40152
CVE-2023-20863	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-20863
CVE-2018-11798	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-17 98

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2019-20330	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-20330
CVE-2019-17531	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-17 531
CVE-2020-11620	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-11620
CVE-2021-31684	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-3 684
CVE-2020-11619	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-11619
CVE-2022-40151	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40151
CVE-2019-17267	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-17 267
CVE-2020-36182	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36182
CVE-2016-6796	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2016-6796
CVE-2019-0205	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-0205
CVE-2018-10054	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-10 054
CVE-2018-14720	MÉDIO	https://nvd.nist. gov/vuln/detail/CVE-2018-14 720
CVE-2022-24329	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2022-24 329

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-36183	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-36183
CVE-2023-44981	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2022-37866	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-3786
CVE-2015-1832	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2015-1832
CVE-2023-38647	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-38647
CVE-2018-1313	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2018-13 13
CVE-2018-12023	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-12 023
CVE-2021-20190	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-2019 0
CVE-2024-38808	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-38808
CVE-2019-20444	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-20 44
GHSA-CHFM-68VV-PvW5	HIGH (ALTO)	https://github.com/advisories/GHSA-CHFM-68VV-PvW5
CVE-2012-0881	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2012-08 81
CVE-2020-36187	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36187

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-13956	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-13956
CVE-2022-38749	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-38749
CVE-2021-37136	NÃO TRIADO	https://nvd.nist. gov/vuln/detail/CVE-2021-37136
CVE-2022-42889	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-4289
CVE-2021-36374	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-36374
CVE-2020-13955	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-1395
CVE-2022-25857	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-25 857

Detalhes dos componentes da versão 7.4.0

Consulte a tabela a seguir para obter mais informações sobre os pacotes extras no Amazon EMR 7.4.0.

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
AmazonCloudWatchAgent	Serviços	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.5-1.amzn2023
AmazonCloudWatchAgent	Serviços	ranger-kms	Extras	2.0.0-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
AmazonCloudWatchAgent	Serviços	s3-dist-cp	Extras	2.33.0-1.amzn2023
AmazonCloudWatchAgent	Serviços	zookeeper	Extras	3.9.1.amzn.1-1.amzn2023
AmazonCloudWatchAgent	Serviços	zookeeper-server	Extras	3.9.1.amzn.1-1.amzn2023
Flink	Serviços	hadoop-client	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-hdfs	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-hdfs-datanode	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-hdfs-namenode	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-httpfs	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-kms	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-mapreduce	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.5-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Flink	Serviços	hadoop-yarn	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-yarn-nodemanager	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-yarn-resourcemanager	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-yarn-timelineserver	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hudi	Extras	0.15.0.amzn.0-1.amzn2023
Flink	Serviços	ranger-kms	Extras	2.0.0-1.amzn2023
Flink	Serviços	hadoop	Extras	3.3.6.amzn.5-1.amzn2023
HBase	Serviços	s3-dist-cp	Extras	2.33.0-1.amzn2023
HCatalog	Serviços	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023
HCatalog	Serviços	mariadb-connector-java	Extras	

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Hive	Serviços	hadoop-httpfs	Extras	3.3.6.amzn.5-1.amzn2023
Hive	Serviços	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023
Hive	Serviços	hive-hbase	Extras	3.1.3.amzn.12-1.amzn2023
Hive	Serviços	mariadb-connector-java	Extras	2.7.2-1
Hive	Serviços	s3-dist-cp	Extras	2.33.0-1.amzn2023
Phoenix	Serviços	s3-dist-cp	Extras	2.33.0-1.amzn2023
Spark	Serviços	spark-datanucleus	Extras	3.5.1.amzn.0-1.amzn2023
Tez	Serviços	hadoop-mapreduce	Extras	3.3.6.amzn.5-1.amzn2023
Tez	Serviços	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.5-1.amzn2023
Tez	Serviços	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	bigtop-utils	Extras	1.2.0-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Trino	Serviços	hadoop-client	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-hdfs	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-hdfs-datanode	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-hdfs-namenode	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-kms	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-mapreduce	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-yarn	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-yarn-nodemanager	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Trino	Serviços	hadoop-yarn-resourcemanager	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-yarn-timelineserver	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hive	Extras	3.1.3.amzn.12-1.amzn2023
Trino	Serviços	hive-hcatalog-server	Extras	3.1.3.amzn.12-1.amzn2023
Trino	Serviços	mariadb-connector-java	Extras	2.7.2-1
Trino	Serviços	ranger-kms	Extras	2.0.0-1.amzn2023
Trino	Serviços	zookeeper	Extras	3.9.1.amzn.1-1.amzn2023
Trino	Serviços	zookeeper-server	Extras	3.9.1.amzn.1-1.amzn2023
Trino	Serviços	bigtop-groovy	Extras	2.5.4-1.amzn2023
Trino	Serviços	bigtop-jsvc	Extras	1.2.4-1.amzn2023
Trino	Serviços	hadoop	Extras	3.3.6.amzn.5-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Trino	Serviços	hadoop-lzo	Extras	0.4.19-1.amzn2023
Trino	Serviços	hive-hcatalog	Extras	3.1.3.amzn.12-1.amzn2023
Trino	Serviços	hive-jdbc	Extras	3.1.3.amzn.12-1.amzn2023
Zookeeper	Serviços	emrfs	Extras	2.64.0-1.amzn2023
Zookeeper	Serviços	hadoop-client	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-hdfs	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-hdfs-datanode	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-hdfs-namenode	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-httpfs	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-kms	Extras	3.3.6.amzn.5-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Zookeeper	Serviços	hadoop-mapreduce	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-yarn	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-yarn-nodemanager	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-yarn-resourcemanager	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-yarn-timelineserver	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	ranger-kms	Extras	2.0.0-1.amzn2023

Notas de lançamento da 7.4.0

As notas de lançamento a seguir incluem informações sobre a versão 7.4.0 do Amazon EMR.

Novos recursos

- Atualizações da aplicação: Amazon EMR 7.4.0 application upgrades include AWS SDK for Java 2.28.8, Delta 3.2.0-amzn-1, Flink 1.19.1, Hadoop 3.4.0, HBase 2.5.5, Hudi 0.15.0-amzn-1, Iceberg 1.6.1-amzn-0, Phoenix 5.2.0, Presto 0.287, Spark 3.5.2, Trino 446, and Zookeeper 3.9.2.

- Esta versão adiciona 7 endpoints de código aberto que oferecem suporte à criptografia em trânsito pela rede. Essa melhoria inclui estruturas de código aberto, como Hue, Livy JupyterEnterpriseGateway, Zookeeper e Ranger KMS. Se você habilitou a criptografia em trânsito e deseja se conectar a esses endpoints, talvez seja necessário modificar seu código para oferecer suporte à comunicação com TLS. Você pode desabilitar o TLS para qualquer um desses endpoints, conforme necessário. Para obter uma lista completa dos endpoints compatíveis, consulte Understanding in-transit encryption.

Problemas conhecidos

- A `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data` propriedade do Yarn ResourceManager, que anteriormente suportava manipuladores aleatórios do Hive desde o emr-6.11.0, agora também aplica o descomissionamento normal aos arquivos servidos pelo serviço de reprodução aleatória externo do Spark, ao executar o Spark no Yarn no emr-7.4.0.

  Observe que o comportamento padrão do serviço de reprodução aleatória externa do Spark, ao executar o Spark no Yarn, é que o Yarn remova os arquivos aleatórios NodeManager do aplicativo quando o aplicativo for encerrado. Isso pode ter um impacto na velocidade de descomissionamento dos nós e na utilização da computação. Para aplicativos de longa execução, considere configurar `spark.shuffle.service.removeShuffle true` para remover arquivos aleatórios que não estão mais em uso, o que permite o descomissionamento mais rápido dos nós sem dados de reprodução aleatória ativos.
- Disponibilidade do Amazon S3 Select — O Amazon S3 Select não oferece suporte a consultas entre regiões para Presto e Trino. Para obter mais informações sobre recursos específicos disponíveis para as versões Presto e Trino, consulte Presto e Trino.
- O Apache Ranger não é compatível com a versão 7.4 do Amazon EMR.
- O Apache Sqoop não é compatível com o Apache Hadoop 3.4 devido a um problema conhecido SQOOP-3485. A versão 7.4 do Amazon EMR inclui, mas não é compatível com o Apache Sqoop. O Apache Sqoop será removido do Amazon EMR versão 7.5 e versões superiores.
- Quando o escalonamento gerenciado está ativado, o YARN ResourceManager (RM) enfrenta um problema crítico de impasse, fazendo com que ele pare de responder ao passar do estado DESCOMISSIONADO para DESCOMISSIONADO com operações simultâneas. Isso acontece principalmente em ambientes de grandes clusters. Quando isso ocorre, tanto as transições de estado do nó quanto as transições de estado do aplicativo falham no processamento, paralisando

efetivamente a funcionalidade principal do. ResourceManager Como solução alternativa, reinicie o RM JVM usando o seguinte comando no nó primário:

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

- Quando um cluster do EMR é iniciado com agendadores YARN diferentes do `CapacityScheduler` ou `FairScheduler` (por exemplo, o Agendador FIFO), a `YARNMemoryAvailablePercentage` métrica não é enviada para. CloudWatch Esse problema afeta a redução de escala em clusters habilitados para escalabilidade gerenciada. Também pode afetar o escalonamento gerenciado personalizado se depender da `YARNMemoryAvailablePercentage` métrica. Como solução alternativa, use o `CapacityScheduler` ou`FairScheduler`.

Alterações, melhorias e problemas resolvidos

- O Amazon EMR usa diferentes versões da TensorFlow biblioteca, dependendo dos tipos de instância que você escolher para o seu cluster. O Amazon EMR também oferece suporte TensorFlow para clusters com tipos de instância aarch64 (instâncias Graviton) do EMR-7.5.0 e superior.
- Otimizando o escalonamento de solicitações do S3 para cargas de trabalho do Iceberg — Para clientes que usam o Iceberg junto com as versões do EMR 34 v7.4, o S3 atualizou o esquema de entropia padrão para um hash de base 2 de 20 caracteres (por exemplo, 35 01010110100110110010). Esse novo esquema de entropia de base 2 reduz o intervalo de caracteres e permite que o S3 escale automaticamente a capacidade de solicitação com mais rapidez para atender às demandas da carga de trabalho alvo do Iceberg. Também reduz a quantidade de tempo em que uma carga de trabalho observa 503 respostas do acelerador.
-

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250527.1	6.1.134-152.225.amzn2023	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein), China (Pequim), Ásia-

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Pacífico (Malásia), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Malásia), Europa (Londres), Ásia-Pacífico (Melbourne), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (AWS GovCloud AWS GovCloud Oregon), México (Central) , Oeste dos EUA (Norte da Califórnia), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv), Ásia-Pacífico (Taipei), Canadá (Central), Oeste

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250512.(	6.1.134-152.225.am zn2023	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Hyderabad), Ásia-Pacífico (Jacarta)), Ásia-Pacífico (Malásia), Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Oriente Médio (Oriente Médio) (Oriente Médio (Oriente Médio)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Emirados Árabes Unidos), Canadá (Central) , Oeste do Canadá (Calgary), (Leste dos EUA), AWS GovCloud (AWS GovCloud Oeste dos EUA), China (Pequim), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250428.1	6.1.134-150.224.amzn2023	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Jacarta) Arta), Ásia-Pacífico (Malásia), Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Malásia) , Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (EAU), Canadá (Central) , Oeste do Canadá

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Calgary), (Leste dos EUA), AWS GovCloud (Oeste dos EUA), China (AWS GovCloud Pequim), China (Ningxia), Israel (Tel Aviv), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250414.(	6.1.132-147.221.am zn2023	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250331.(	6.1.131-143.221.am zn2023	18 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central), Ásia-Pacífico (Tóquio)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250303.(	6.1.129-138.220.am zn2023	27 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250218.1	6.1.128-136.201.amzn2023	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderabad), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250211.(	6.1.127-135.201.am zn2023	26 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250123.4	6.1.124-134.200.amzn2023	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderabad), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250115.(	6.1.119-129.201.am zn2023	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 241121.(	6.1.115-126.197.am zn2023	12 de dezembro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.5.2 240916	6.1.109-118.189.am zn2023	6 de setembro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

Versões Java padrão 7.4.0

O Amazon EMR versão 7.4 e superior vem com o Amazon Corretto 17 (construído no OpenJDK) por padrão para aplicativos que suportam o Corretto 17 (JDK 17), com exceção do Apache Livy.

A tabela a seguir mostra as versões padrão do Java para aplicativos no Amazon EMR 7.4.0. Se você quiser alterar a JVM padrão no cluster, siga as instruções em Configurar aplicações para usar uma Máquina Virtual Java específica para cada aplicação executada no cluster. Você só pode usar uma versão de runtime do Java para um cluster. O Amazon EMR não é compatível com a execução de diferentes nós ou aplicações em diferentes versões de runtime no mesmo cluster.

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
Delta	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez	17, 11, 8
Trino	21, 17
Zeppelin	17, 11, 8
Pig	8
Zookeeper	8

Versões do componente 7.4.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`*

começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
adot-java-agent	1.31.0	Um Java Agent que coleta métricas de daemons de aplicações.
delta	3.2.0-amzn-1	O Delta Lake é um formato de tabela aberto para grandes conjuntos de dados analíticos
emr-amazon-cloudwatch-agent	1.300032.2-amzn-0	Um aplicativo que coleta métricas internas em nível de sistema e métricas de aplicativos personalizadas de instâncias da Amazon. EC2
emr-ddb	5.6.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.13.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.17.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.17.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.34.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.

Componente	Versão	Descrição
emr-s3-select	2.13.0	EMR S3Select Connector
emr-wal-cli	1.3.1	Cli usada para lista/exclusão de emrwal.
emrfs	2.65,0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.19.1-amzn-0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.19.1-amzn-0	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
hadoop-client	3.4.0-amzn-0	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	3.4.0-amzn-0	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.4.0-amzn-0	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.4.0-amzn-0	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	3.4.0-amzn-0	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.

Componente	Versão	Descrição
hadoop-httpfs-server	3.4.0-amzn-0	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.4.0-amzn-0	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	3.4.0-amzn-0	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.4.0-amzn-0	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	3.4.0-amzn-0	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	3.4.0-amzn-0	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	2.5.5-amzn-0	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	2.5.5-amzn-0	Serviço para atender a uma ou mais HBase regiões.
hbase-client	2.5.5-amzn-0	HBase cliente de linha de comando.

Componente	Versão	Descrição
hbase-rest-server	2.5.5-amzn-0	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.5.5-amzn-0	Serviço que fornece um endpoint Thrift para. HBase
hbase-operator-tools	2.5.5-amzn-0	Ferramenta de reparo para HBase clusters Apache.
hcatalog-client	3.1.3-amzn-13	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.3-amzn-13	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	3.1.3-amzn-13	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.3-amzn-13	O cliente da linha de comando do Hive.
hive-hbase	3.1.3-amzn-13	Cliente do Hive-hbase.
hive-metastore-server	3.1.3-amzn-13	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.

Componente	Versão	Descrição
hive-server2	3.1.3-amzn-13	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.15.0-amzn-1	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.15.0-amzn-1	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-trino	0.15.0-amzn-1	Biblioteca de pacotes para execução do Presto com o Hudi.
hudi-spark	0.15.0-amzn-1	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.11.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
iceberg	1.6.1-amzn-0	Apache Iceberg é um formato de tabela aberta para conjuntos de dados analíticos imensos
jupyterhub	1.5.0	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.8.0-incubating	Interface REST para interagir com o Apache Spark

Componente	Versão	Descrição
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.
nvidia-cuda	12.3.0	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.7.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	5.2.0	As bibliotecas do phoenix para servidor e cliente
phoenix-connectors	6.0.0	Conectores do Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0,287-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.

Componente	Versão	Descrição
presto-worker	0,287-amzn-0	O serviço que executa partes de uma consulta.
presto-client	0,287-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
trino-coordinator	446-aman-0	Serviço que aceita consultas e gerencia a execução de consultas entre trino-workers.
trino-worker	446-aman-0	O serviço que executa partes de uma consulta.
trino-client	446-aman-0	Cliente de linha de comando do Trino instalado em mestres de espera de um cluster de alta disponibilidade (HA) em que o servidor do Trino não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.3.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	3.5.2-amzn-0	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	3.5.2-amzn-0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.5.2-amzn-0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.5.2-amzn-0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
spark-rapids	24.08.1-amzn-0	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.16.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.10.2-amzn-11	O aplicativo e as bibliotecas do YARN para Tez.
tez-on-worker	0.10.2-amzn-11	A aplicação e as bibliotecas do YARN nós de processam ento.
zeppelin-server	0.11.1	O notebook baseado na web que permite um data analytics interativo.

Componente	Versão	Descrição
zookeeper-server	3.9.2-amzn-0	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.9.2-amzn-0	ZooKeeper cliente de linha de comando.

Classificações de configuração 7.4.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

classificações do emr-7.4.0

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Classificações	Descrição	Ações de reconfiguração
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.

Classificações	Descrição	Ações de reconfiguração
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Alterar a página inicial do KMS do Hadoop	Not available.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.

Classificações	Descrição	Ações de reconfiguração
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.	Not available.
iceberg-defaults	Alterar os valores no arquivo iceberg-defaults.conf do Iceberg.	Not available.
delta-defaults	Alterar os valores no arquivo delta-defaults.conf do Delta.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j2	Alterar as configurações de log4j2.properties do Livy.	Restarts Livy Server.

Classificações	Descrição	Ações de reconfiguração
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j2	Altere os valores no arquivo log4j2.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.

Classificações	Descrição	Ações de reconfiguração
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-lakeformation	Alterar os valores no arquivo lakeformation.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
trino-log	Alterar os valores no arquivo log.properties do Trino.	Restarts Trino-Server (for Trino)
trino-config	Alterar os valores no arquivo config.properties do Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Alterar os valores no arquivo password-authenticator.properties do Trino.	Restarts Trino-Server (for Trino)
trino-env	Alterar os valores no arquivo trino-env.sh do Trino.	Restarts Trino-Server (for Trino)
trino-node	Alterar os valores no arquivo node.properties do Trino.	Not available.
trino-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Trino.	Not available.
trino-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Trino.	Not available.
trino-connector-delta	Alterar os valores no arquivo delta.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Alterar os valores no arquivo hive.properties do Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Alterar os valores no arquivo exchange-manager.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Alterar os valores no arquivo iceberg.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Alterar os valores no arquivo hudi.properties do Trino.	Restarts Trino-Server (for Trino)

Classificações	Descrição	Ações de reconfiguração
trino-connector-jmx	Alterar os valores no arquivo jmx.properties do Trino.	Not available.
trino-connector-kafka	Alterar os valores no arquivo kafka.properties do Trino.	Not available.
trino-connector-localfile	Alterar os valores no arquivo localfile.properties do Trino.	Not available.
trino-connector-memory	Alterar os valores no arquivo memory.properties do Trino.	Not available.
trino-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Trino.	Not available.
trino-connector-mysql	Alterar os valores no arquivo mysql.properties do Trino.	Not available.
trino-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Trino.	Not available.
trino-connector-raptor	Alterar os valores no arquivo raptor.properties do Trino.	Not available.
trino-connector-redis	Alterar os valores no arquivo redis.properties do Trino.	Not available.
trino-connector-redshift	Alterar os valores no arquivo redshift.properties do Trino.	Not available.
trino-connector-tpch	Alterar os valores no arquivo tpch.properties do Trino.	Not available.
trino-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Trino.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Alterar os valores no arquivo kms-logback.xml do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j2	Alterar os valores no arquivo log4j2.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie and HiveServe r2.
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.

Classificações	Descrição	Ações de reconfiguração
zookeeper-logback	Altere os valores no ZooKeeper arquivo logback.xml do.	Restarts Zookeeper server.
emr-metrics	Altere as configurações de métricas do EMR desse nó.	Restarts the CloudWatchAgent service.

Registro de alterações 7.4.0

Registro de alterações da versão 7.4.0 e das notas de lançamento

Data	Event	Descrição
18/11/2024	Publicação de documentos	Notas de lançamento do Amazon EMR 7.4.0 publicadas pela primeira vez
2024-11-13	Implantação concluída	Amazon EMR 7.4.0 totalmente implantado em todas as regiões suportadas
31/10/2024	Início da implantação	Amazon EMR 7.4.0 implantado pela primeira vez nas regiões comerciais iniciais

Versão 7.3.0 do Amazon EMR

Ciclo de vida compatível com a versão 7.3.0

A tabela a seguir descreve as datas de ciclo de vida compatíveis com a versão 7.3.0 do Amazon EMR.

Fase de suporte	Data	
Data da versão inicial	16 de outubro de 2024	

Fase de suporte	Data
Suporte padrão até	15 de outubro de 2026
Término do suporte	16 de outubro de 2026
Fim da vida útil	16 de outubro de 2027

Para obter uma lista de vulnerabilidades e exposições (CVEs) comuns críticas e altas que não afetam os clusters do EMR de acordo com o software e as configurações recomendados, consulte 7.3.0 vulnerabilidades e exposições comuns conhecidas para mecanismos principais.

Componentes da versão 7.3.0

O Amazon EMR 7.3.0 inclui os seguintes aplicativos:

Mecanismos centrais:

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

Extras:

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

Para obter mais detalhes sobre os pacotes incluídos, consulte Versões do aplicativo nas versões 7.x do Amazon EMR. Para obter mais informações sobre os mecanismos centrais e extras, consulte Amazon EMR Standard Support.

Versões 7.3.0 da aplicação

Essa versão inclui os seguintes aplicativos: AmazonCloudWatchAgent, Delta, Flink, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, Oozie, Phoenix, Pig, Presto, Spark, SqoopTensorFlow, Tez, Trino, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
AWS SDK for Java	2.25.70, 1.12.747	2.23.18, 1.12.705	2.23.18, 1.12.656	2.20.160-amzn-0, 1.12.569
Python	3,9, 3,11	3,9, 3,11	3,9, 3,11	3.9
Scala	2.12.18	2.12.17	2.12.17	2.12.17
AmazonCloudWatchAgent	1.300032.2	1.300032.2	1.300032.2	1.300031.1
Delta	3.2.0-amzn-0	3.1.0	3.0.0	3.0.0
Flink	1.18.1	1.18.1	1.18.1	1.18.0
Ganglia	-	-	-	-
HBase	2.4.17	2.4.17	2.4.17	2.4.17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.6	3.3.6	3.3.6	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-amzn-0	0.14.1-amzn-1	0.14.1-amzn-0	0.14.0-amzn-1
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.5.2-amzn-0	1.5.0-amzn-0	1.4.3-amzn-0	q1.4.2-amzn-0

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.7.1
MXNet	-	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,285	0,285	0,284	0,283
Spark	3.5.1	3.5.1	3.5.0	3.5.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.16.1	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	442	436	435	426
Zeppelin	0.11.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.9.1	3.9.1	3.9.1	3.5.10

Vulnerabilidades e exposições comuns da versão 7.3.0

A tabela a seguir lista todos os CVEs que não afetam os clusters do EMR que são executados nas configurações recomendadas do Amazon EMR 7.3.0. O Amazon EMR depende do código upstream aberto para disponibilidade de correções e fornecerá a versão estável mais recente como parte dos

componentes do mecanismo central em até 90 dias após a verificação das correções pelo Amazon EMR.

Essa tabela foi atualizada em 4 de outubro de 2024.

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2018-1000180	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-1000180
CVE-2020-28052	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-28052
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2021-31684	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-3 684
CVE-2022-1471	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-1471
CVE-2022-25857	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-25 857
CVE-2022-3171	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3510
CVE-2022-40149	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40149
CVE-2022-40150	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40150

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2022-40152	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40152
CVE-2022-42003	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2003
CVE-2022-42003	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2003
CVE-2022-42004	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2004
CVE-2022-45685	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-45 685
CVE-2022-45693	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-45693
CVE-2023-1370	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2023-1436	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-36478	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-36478
CVE-2023-39410	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-44981	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-44981

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2023-50387	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-50387
CVE-2012-0881	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2012-08 81
CVE-2013-4002	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2013-4002
CVE-2015-1832	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2015-1832
CVE-2015-5237	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2015-5237
CVE-2018-11307	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-11 307
CVE-2018-12022	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2018-12022
CVE-2018-12023	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-12 023
CVE-2018-14718	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14 718
CVE-2018-14719	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14 719
CVE-2018-14720	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14 720
CVE-2018-14721	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14 721
CVE-2018-19360	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-19360

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2018-19361	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-1936 1
CVE-2018-19362	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-1936 2
CVE-2018-7489	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-7 489
CVE-2019-12086	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-12 2086
CVE-2019-14379	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14 379
CVE-2019-14439	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-14 439
CVE-2019-14540	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14 540
CVE-2019-14892	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14 892
CVE-2019-14893	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14 893
CVE-2019-16335	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16 335
CVE-2019-16942	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16 942
CVE-2019-16943	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16 943
CVE-2019-17267	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17 267

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2019-17531	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17 531
CVE-2019-20330	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20330
CVE-2020-10650	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10650
CVE-2020-10672	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10 672
CVE-2020-10673	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10 673
CVE-2020-10968	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10 968
CVE-2020-10969	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10 969
CVE-2020-11111	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-1111
CVE-2020-11112	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-1112
CVE-2020-11113	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-11-13
CVE-2020-11619	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-11619
CVE-2020-11620	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-11620
CVE-2020-13936	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-13936

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-13949	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2020-14060	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14060
CVE-2020-14061	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14061
CVE-2020-14062	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14062
CVE-2020-14195	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14 195
CVE-2020-24616	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-24 616
CVE-2020-24750	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-24750
CVE-2020-25649	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-25 649
CVE-2020-35490	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-35490
CVE-2020-35491	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-35491
CVE-2020-35728	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-35728
CVE-2020-36179	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36179
CVE-2020-36180	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36180

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-36181	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36181
CVE-2020-36182	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36182
CVE-2020-36183	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36183
CVE-2020-36184	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36184
CVE-2020-36185	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36185
CVE-2020-36186	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36186
CVE-2020-36187	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36187
CVE-2020-36188	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-3618
CVE-2020-36189	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36189
CVE-2020-36518	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36518
CVE-2020-8840	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-8840
CVE-2020-9546	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9546
CVE-2020-9547	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9547

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-9548	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9548
CVE-2021-20190	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-2019 0
CVE-2021-31684	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-3 684
CVE-2021-34538	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-34 538
CVE-2021-34538	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-34 538
CVE-2021-35515	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-35515
CVE-2021-35516	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-35516
CVE-2021-35517	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-35517
CVE-2021-36090	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-36090
CVE-2022-25647	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-25 647
CVE-2022-3171	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3510

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2022-36364	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3636 4
CVE-2022-37865	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-37865
CVE-2022-37866	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3786
CVE-2022-39135	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-39135
CVE-2022-42003	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2003
CVE-2022-42004	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2004
CVE-2022-46337	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-46337
CVE-2022-46751	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-46751
CVE-2023-1370	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-34453	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3445 3

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2023-34454	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3445
CVE-2023-34455	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3445
CVE-2023-39410	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-43642	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2016-5018	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2016-5018
CVE-2016-6796	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2016-6796
CVE-2019-0205	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-0205
CVE-2020-13949	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-43642	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2015-2156	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2015-2156
CVE-2017-18640	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2017-18 640
CVE-2017-18640	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2017-18 640

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2019-16869	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-16 869
CVE-2019-20444	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20 44
CVE-2019-20445	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20 445
CVE-2021-37136	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-37136
CVE-2021-37137	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-37137
CVE-2022-1471	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-1471
CVE-2022-1471	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-1471
CVE-2022-25857	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-25 857
CVE-2022-25857	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-25 857
CVE-2022-3171	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3510
CVE-2022-40149	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40149

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2022-40150	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40150
CVE-2022-42003	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2003
CVE-2022-42004	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2004
CVE-2022-45685	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-45 685
CVE-2022-45693	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-45693
CVE-2023-1436	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-36478	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-36478
CVE-2023-44981	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2023-44981	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2019-0205	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-0205
CVE-2020-13949	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-13949

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2021-34538	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-34 538
CVE-2022-45688	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-4568
CVE-2022-46337	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-46337
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-3635	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3635
CVE-2023-5072	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-5072
CVE-2024-21634	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-21634
CVE-2022-40149	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40149
CVE-2022-40150	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40150
CVE-2022-45685	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-45 685
CVE-2022-45693	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-45693
CVE-2023-1436	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2023-6378	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-6378

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2023-6378	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-6378
CVE-2023-6481	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-6481
CVE-2023-6481	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-6481

Detalhes dos componentes da versão 7.3.0

Consulte a tabela a seguir para obter mais informações sobre os pacotes Extras no Amazon EMR 7.3.0.

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
AmazonCloudWatchAgent	Serviços	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.5-1.amzn2023
AmazonCloudWatchAgent	Serviços	ranger-kms	Extras	2.0.0-1.amzn2023
AmazonCloudWatchAgent	Serviços	s3-dist-cp	Extras	2.33.0-1.amzn2023
AmazonCloudWatchAgent	Serviços	zookeeper	Extras	3.9.1.amzn.1-1.amzn2023
AmazonCloudWatchAgent	Serviços	zookeeper-server	Extras	3.9.1.amzn.1-1.amzn2023
Flink	Serviços	hadoop-client	Extras	3.3.6.amzn.5-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Flink	Serviços	hadoop-hdfs	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-hdfs-datanode	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-hdfs-namenode	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-httpfs	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-kms	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-mapreduce	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-yarn	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-yarn-nodemanager	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Flink	Serviços	hadoop-yarn-resourcemanager	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hadoop-yarn-timelineserver	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Serviços	hudi	Extras	0.15.0.amzn.0-1.amzn2023
Flink	Serviços	ranger-kms	Extras	2.0.0-1.amzn2023
Flink	Serviços	hadoop	Extras	3.3.6.amzn.5-1.amzn2023
HBase	Serviços	s3-dist-cp	Extras	2.33.0-1.amzn2023
HCatalog	Serviços	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023
HCatalog	Serviços	mariadb-connector-java	Extras	
Hive	Serviços	hadoop-httpfs	Extras	3.3.6.amzn.5-1.amzn2023
Hive	Serviços	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023
Hive	Serviços	hive-hbase	Extras	3.1.3.amzn.12-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Hive	Serviços	mariadb-connector-java	Extras	2.7.2-1
Hive	Serviços	s3-dist-cp	Extras	2.33.0-1.amzn2023
Phoenix	Serviços	s3-dist-cp	Extras	2.33.0-1.amzn2023
Spark	Serviços	spark-datanucleus	Extras	3.5.1.amzn.0-1.amzn2023
Tez	Serviços	hadoop-mapreduce	Extras	3.3.6.amzn.5-1.amzn2023
Tez	Serviços	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.5-1.amzn2023
Tez	Serviços	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	bigtop-utils	Extras	1.2.0-1.amzn2023
Trino	Serviços	hadoop-client	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-hdfs	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-hdfs-datanode	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.5-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Trino	Serviços	hadoop-hdfs-namenode	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-kms	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-mapreduce	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-yarn	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-yarn-nodemanager	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-yarn-resourcemanager	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-yarn-timelineserver	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hive	Extras	3.1.3.amzn.12-1.amzn2023
Trino	Serviços	hive-hcatalog-server	Extras	3.1.3.amzn.12-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Trino	Serviços	mariadb-connector-java	Extras	2.7.2-1
Trino	Serviços	ranger-kms	Extras	2.0.0-1.amzn2023
Trino	Serviços	zookeeper	Extras	3.9.1.amzn.1-1.amzn2023
Trino	Serviços	zookeeper-server	Extras	3.9.1.amzn.1-1.amzn2023
Trino	Serviços	bigtop-groovy	Extras	2.5.4-1.amzn2023
Trino	Serviços	bigtop-jsvc	Extras	1.2.4-1.amzn2023
Trino	Serviços	hadoop	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Serviços	hadoop-lzo	Extras	0.4.19-1.amzn2023
Trino	Serviços	hive-hcatalog	Extras	3.1.3.amzn.12-1.amzn2023
Trino	Serviços	hive-jdbc	Extras	3.1.3.amzn.12-1.amzn2023
Zookeeper	Serviços	emrfs	Extras	2.64.0-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Zookeeper	Serviços	hadoop-client	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-hdfs	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-hdfs-datanode	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-hdfs-namenode	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-httpfs	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-kms	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-mapreduce	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-yarn	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-yarn-nodemanager	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Zookeeper	Serviços	hadoop-yarn-resourcemanager	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	hadoop-yarn-timelineserver	Extras	3.3.6.amzn.5-1.amzn2023
Zookeeper	Serviços	ranger-kms	Extras	2.0.0-1.amzn2023

Notas da versão 7.3.0

As notas da versão a seguir incluem informações sobre a versão 7.3.0 do Amazon EMR. As alterações são referentes à versão 7.3.0.

Novos recursos

- Atualizações da aplicação: Amazon EMR 7.3.0 application upgrades include Delta 3.2.0-amzn-0, Hudi 0.15.0-amzn-0, TensorFlow 2.16.1, Trino 442, and Zeppelin 0.11.1.
- Esta versão adiciona 22 endpoints de código aberto que oferecem suporte à criptografia em trânsito pela rede. Essa melhoria inclui estruturas de código aberto, como Hadoop, Hive, Tez, HBase Phoenix e Flink. Se você habilitou a criptografia em trânsito e deseja se conectar a esses endpoints, talvez seja necessário modificar seu código para oferecer suporte à comunicação com TLS. Você pode desabilitar o TLS para qualquer um desses endpoints, conforme necessário. Para obter uma lista completa dos endpoints compatíveis, consulte Understanding in-transit encryption.
- Com o Amazon EMR 7.3.0, o Apache Hive com Iceberg agora oferece suporte ao AWS Glue Data Catalog como metastore, o que oferece benefícios como um repositório centralizado para armazenar metadados, rastrear automaticamente fontes de dados e extrair metadados, além de integração com outros serviços para fluxos de trabalho contínuos.

Problemas conhecidos

- Não há suporte para o Python 3.11 no EMR Studio.

- Se você usar criptografia em trânsito e autenticação do Kerberos, não poderá usar aplicações do Spark que dependam do Hive Metastore. Esse problema foi corrigido no HIVE-16340 e será totalmente resolvido no SPARK-44114 quando o Spark de código aberto puder ser atualizado para o Hive 3. Enquanto isso, você pode contornar temporariamente esse problema configurando `hive.metastore.use.SSL` como `false`. Para obter mais informações, consulte Configure applications.
  - Se você não usa a autenticação do Kerberos, deve configurar o Spark para usar SSL. Para isso, defina as configurações a seguir ao provisionar o cluster.

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

- Se você usa domínios personalizados, alguns links na interface do usuário do Spark History Server no AWS console podem não funcionar. Faça upload do script a seguir em um bucket do S3 e, em seguida, adicione uma etapa para usar o script de shell.

```
#!/bin/bash

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

sudo systemctl restart nginx
```

Alterações, melhorias e problemas resolvidos

- Para oferecer suporte à verificação do nome de host TLS para o YARN, essa versão altera a configuração `yarn.resourcemanager.hostname` do endereço IP da instância primária para o nome DNS privado da instância primária.
- MXNet o projeto chegou ao fim da vida útil e agora está arquivado. the section called “emr-7.2.0”é a última versão que inclui MXNet.

- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250527.1	6.1.134-152.225.am zn2023	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Malásia), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Malásia), Europa (Londres), Ásia-Pacífico (Melbourne), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (AWS GovCloud AWS GovCloud Oregon), México (Central) , Oeste dos EUA (Norte da Califórnia), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv), Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250512.(	6.1.134-152.225.am zn2023	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Hyderabad), Ásia-Pacífico (Jacarta)), Ásia-Pacífico (Malásia), Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Oriente Médio (Oriente Médio) (Oriente Médio (Oriente Médio)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Emirados Árabes Unidos), Canadá (Central) , Oeste do Canadá (Calgary), (Leste dos EUA), AWS GovCloud (AWS GovCloud Oeste dos EUA), China (Pequim), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250428.1	6.1.134-150.224.amzn2023	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Jacarta) Arta), Ásia-Pacífico (Malásia), Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Malásia) , Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (EAU), Canadá (Central) , Oeste do Canadá

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Calgary), (Leste dos EUA), AWS GovCloud (Oeste dos EUA), China (AWS GovCloud Pequim), China (Ningxia), Israel (Tel Aviv), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250414.(	6.1.132-147.221.am zn2023	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250331.(	6.1.131-143.221.am zn2023	18 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central), Ásia-Pacífico (Tóquio)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250303.(	6.1.129-138.220.am zn2023	27 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250218.1	6.1.128-136.201.amzn2023	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderabad), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250211.(	6.1.127-135.201.am zn2023	26 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250123.4	6.1.124-134.200.amzn2023	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderabad), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250115.(	6.1.119-129.201.am zn2023	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 241121.(	6.1.115-126.197.am zn2023	12 de dezembro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 241031.(	6.1.112-124.190.am zn2023	15 de novembro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.5.2 240916	6.1.109-118.189.am zn2023	6 de setembro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

Versões padrão do Java 7.3.0

O Amazon EMR 7.3 e versões superiores vêm com o Amazon Corretto 17 (criado no OpenJDK) por padrão para aplicações compatíveis com o Corretto 17 (JDK 17), com exceção do Apache Livy.

A tabela a seguir mostra as versões padrão do Java para aplicações no Amazon EMR 7.3.0. Se você quiser alterar a JVM padrão no cluster, siga as instruções em Configurar aplicações para usar uma Máquina Virtual Java específica para cada aplicação executada no cluster. Você só pode usar uma versão de runtime do Java para um cluster. O Amazon EMR não é compatível com a execução de diferentes nós ou aplicações em diferentes versões de runtime no mesmo cluster.

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
Delta	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez	17, 11, 8
Trino	21, 17
Zeppelin	8
Pig	8
Zookeeper	17, 11, 8

Versões de componente 7.3.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`*

começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
adot-java-agent	1.31.0	Um Java Agent que coleta métricas de daemons de aplicações.
delta	3.2.0-amzn-0	O Delta Lake é um formato de tabela aberto para grandes conjuntos de dados analíticos
emr-amazon-cloudwatch-agent	1.300032.2-amzn-0	Um aplicativo que coleta métricas internas em nível de sistema e métricas de aplicativos personalizadas de instâncias da Amazon. EC2
emr-ddb	5.5.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.12.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.16.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.16.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.33.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.

Componente	Versão	Descrição
emr-s3-select	2.12.0	EMR S3Select Connector
emr-wal-cli	1.3.0	Cli usada para lista/exclusão de emrwal.
emrfs	2.64.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.18.1-amzn-2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.18.1-amzn-2	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
hadoop-client	3.3.6-amzn-5	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	3.3.6-amzn-5	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.3.6-amzn-5	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.3.6-amzn-5	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	3.3.6-amzn-5	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.

Componente	Versão	Descrição
hadoop-httpfs-server	3.3.6-amzn-5	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.3.6-amzn-5	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	3.3.6-amzn-5	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.3.6-amzn-5	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	3.3.6-amzn-5	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	3.3.6-amzn-5	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	2.4.17-amzn-7	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	2.4.17-amzn-7	Serviço para atender a uma ou mais HBase regiões.
hbase-client	2.4.17-amzn-7	HBase cliente de linha de comando.

Componente	Versão	Descrição
hbase-rest-server	2.4.17-amzn-7	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.4.17-amzn-7	Serviço que fornece um endpoint Thrift para. HBase
hbase-operator-tools	2.4.17-amzn-7	Ferramenta de reparo para HBase clusters Apache.
hcatalog-client	3.1.3-amzn-12	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.3-amzn-12	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	3.1.3-amzn-12	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.3-amzn-12	O cliente da linha de comando do Hive.
hive-hbase	3.1.3-amzn-12	Cliente do Hive-hbase.
hive-metastore-server	3.1.3-amzn-12	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.

Componente	Versão	Descrição
hive-server2	3.1.3-amzn-12	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.15.0-amzn-0	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.15.0-amzn-0	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-trino	0.15.0-amzn-0	Biblioteca de pacotes para execução do Presto com o Hudi.
hudi-spark	0.15.0-amzn-0	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.11.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
iceberg	1.5.2-amzn-0	Apache Iceberg é um formato de tabela aberta para conjuntos de dados analíticos imensos
jupyterhub	1.5.0	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.8.0-incubating	Interface REST para interagir com o Apache Spark

Componente	Versão	Descrição
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.
nvidia-cuda	12.3.0	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.7.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	5.1.3	As bibliotecas do phoenix para servidor e cliente
phoenix-connectors	6.0.0	Conectores do Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.285-amzn-1	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.

Componente	Versão	Descrição
presto-worker	0.285-amzn-1	O serviço que executa partes de uma consulta.
presto-client	0.285-amzn-1	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
trino-coordinator	442-amzn-0	Serviço que aceita consultas e gerencia a execução de consultas entre trino-workers.
trino-worker	442-amzn-0	O serviço que executa partes de uma consulta.
trino-client	442-amzn-0	Cliente de linha de comando do Trino instalado em mestres de espera de um cluster de alta disponibilidade (HA) em que o servidor do Trino não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.3.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	3.5.1-amzn-1	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	3.5.1-amzn-1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.5.1-amzn-1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.5.1-amzn-1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
spark-rapids	24.06.1-amzn-0	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.16.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.10.2-amzn-10	O aplicativo e as bibliotecas do YARN para Tez.
tez-on-worker	0.10.2-amzn-10	A aplicação e as bibliotecas do YARN nós de processam ento.
zeppelin-server	0.11.1	O notebook baseado na web que permite um data analytics interativo.

Componente	Versão	Descrição
zookeeper-server	3.9.1-amzn-1	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.9.1-amzn-1	ZooKeeper cliente de linha de comando.

Classificações de configuração 7.3.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-7.3.0

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Classificações	Descrição	Ações de reconfiguração
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.

Classificações	Descrição	Ações de reconfiguração
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Alterar a página inicial do KMS do Hadoop	Not available.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.

Classificações	Descrição	Ações de reconfiguração
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.	Not available.
iceberg-defaults	Alterar os valores no arquivo iceberg-defaults.conf do Iceberg.	Not available.
delta-defaults	Alterar os valores no arquivo delta-defaults.conf do Delta.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j2	Alterar as configurações de log4j2.properties do Livy.	Restarts Livy Server.

Classificações	Descrição	Ações de reconfiguração
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.

Classificações	Descrição	Ações de reconfiguração
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-lakeformation	Alterar os valores no arquivo lakeformation.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
trino-log	Alterar os valores no arquivo log.properties do Trino.	Restarts Trino-Server (for Trino)
trino-config	Alterar os valores no arquivo config.properties do Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Alterar os valores no arquivo password-authenticator.properties do Trino.	Restarts Trino-Server (for Trino)
trino-env	Alterar os valores no arquivo trino-env.sh do Trino.	Restarts Trino-Server (for Trino)
trino-node	Alterar os valores no arquivo node.properties do Trino.	Not available.
trino-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Trino.	Not available.
trino-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Trino.	Not available.
trino-connector-delta	Alterar os valores no arquivo delta.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Alterar os valores no arquivo hive.properties do Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Alterar os valores no arquivo exchange-manager.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Alterar os valores no arquivo iceberg.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Alterar os valores no arquivo hudi.properties do Trino.	Restarts Trino-Server (for Trino)

Classificações	Descrição	Ações de reconfiguração
trino-connector-jmx	Alterar os valores no arquivo jmx.properties do Trino.	Not available.
trino-connector-kafka	Alterar os valores no arquivo kafka.properties do Trino.	Not available.
trino-connector-localfile	Alterar os valores no arquivo localfile.properties do Trino.	Not available.
trino-connector-memory	Alterar os valores no arquivo memory.properties do Trino.	Not available.
trino-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Trino.	Not available.
trino-connector-mysql	Alterar os valores no arquivo mysql.properties do Trino.	Not available.
trino-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Trino.	Not available.
trino-connector-raptor	Alterar os valores no arquivo raptor.properties do Trino.	Not available.
trino-connector-redis	Alterar os valores no arquivo redis.properties do Trino.	Not available.
trino-connector-redshift	Alterar os valores no arquivo redshift.properties do Trino.	Not available.
trino-connector-tpch	Alterar os valores no arquivo tpch.properties do Trino.	Not available.
trino-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Trino.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Alterar os valores no arquivo kms-logback.xml do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j2	Alterar os valores no arquivo log4j2.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie and HiveServe r2.
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.

Classificações	Descrição	Ações de reconfiguração
zookeeper-logback	Altere os valores no ZooKeeper arquivo logback.xml do.	Restarts Zookeeper server.
emr-metrics	Altere as configurações de métricas do EMR desse nó.	Restarts the CloudWatchAgent service.

Log de alterações 7.3.0

Log de alterações da versão 7.3.0 e das notas da versão

Data	Event	Descrição
2024-10-17	Publicação de documentos	Notas da versão 7.3.0 do Amazon EMR publicadas pela primeira vez
16/10/2024	Implantação concluída	Amazon EMR 7.3.0 totalmente implantado em todas as regiões compatíveis
14/10/2024	Lançamento inicial	A versão 7.3.0 do Amazon EMR foi implantada pela primeira vez em regiões comerciais iniciais

Versão 7.2.0 do Amazon EMR

Ciclo de vida compatível com a versão 7.2.0

A tabela a seguir descreve as datas de ciclo de vida compatíveis na versão 7.2.0 do Amazon EMR.

Fase de suporte	Data	
Data da versão inicial	25 de julho de 2024	

Fase de suporte	Data
Suporte padrão até	24 de julho de 2026
Término do suporte	25 de julho de 2026
Fim da vida útil	25 de julho de 2027

Para obter uma lista de vulnerabilidades e exposições (CVEs) comuns críticas e altas que não afetam os clusters do EMR de acordo com o software e as configurações recomendados, consulte 7.2.0 vulnerabilidades e exposições comuns conhecidas para mecanismos principais. Para ver uma lista dos CVEs que foram corrigidos na versão 7.2.0, consulte 7.2.0 corrigiu vulnerabilidades e exposições comuns.

Componentes da versão 7.2.0

O Amazon EMR 7.2.0 inclui os seguintes aplicativos:

Mecanismos centrais:

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

Extras:

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

Para obter mais detalhes sobre os pacotes incluídos, consulte Versões do aplicativo nas versões 7.x do Amazon EMR. Para obter mais informações sobre os mecanismos centrais e extras, consulte Amazon EMR Standard Support.

Versões 7.2.0 da aplicação

Essa versão inclui os seguintes aplicativos: AmazonCloudWatchAgent, Delta, Flink, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, SqoopTensorFlow, Tez, TrinoZeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
AWS SDK for Java	2.25.70, 1.12.747	2.23.18, 1.12.705	2.23.18, 1.12.656	2.20.160-amzn-0, 1.12.569
Python	3,9, 3,11	3,9, 3,11	3,9, 3,11	3.9
Scala	2.12.18	2.12.17	2.12.17	2.12.17
AmazonCloudWatchAgent	1.300032.2	1.300032.2	1.300032.2	1.300031.1
Delta	3.2.0-amzn-0	3.1.0	3.0.0	3.0.0
Flink	1.18.1	1.18.1	1.18.1	1.18.0
Ganglia	-	-	-	-
HBase	2.4.17	2.4.17	2.4.17	2.4.17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.6	3.3.6	3.3.6	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-amzn-0	0.14.1-amzn-1	0.14.1-amzn-0	0.14.0-amzn-1
Hue	4.11.0	4.11.0	4.11.0	4.11.0

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
Iceberg	1.5.2-amzn-0	1.5.0-amzn-0	1.4.3-amzn-0	q1.4.2-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.7.1
MXNet	-	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,285	0,285	0,284	0,283
Spark	3.5.1	3.5.1	3.5.0	3.5.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.16.1	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	442	436	435	426
Zeppelin	0.11.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.9.1	3.9.1	3.9.1	3.5.10

Vulnerabilidades e exposições comuns da versão 7.2.0

Esta página lista todas as vulnerabilidades e exposições comuns (CVEs) relacionadas ao Amazon EMR 7.2.0.

Vulnerabilidades e exposições comuns conhecidas em mecanismos centrais da versão 7.2.0

A tabela a seguir lista todos os CVEs que não afetam os clusters do EMR que são executados nas configurações recomendadas do Amazon EMR 7.2.0. O Amazon EMR depende do código aberto upstream para disponibilidade de correções e fornecerá a versão estável mais recente como parte dos componentes do mecanismo central em até 90 dias após a verificação das correções pelo Amazon EMR.

Esta tabela foi atualizada em 16 de setembro de 2024.

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2021-31684	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-3 684
CVE-2022-1471	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-1471
CVE-2022-25857	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-25 857
CVE-2022-3171	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3510
CVE-2022-40149	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40149

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2022-40150	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40150
CVE-2022-40152	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40152
CVE-2022-42003	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2003
CVE-2022-42003	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2003
CVE-2022-42004	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2004
CVE-2022-45685	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-45 685
CVE-2022-45693	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-45693
CVE-2023-1370	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2023-1436	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-36478	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-36478
CVE-2023-39410	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-39410

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2023-44981	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2015-5237	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2015-5237
CVE-2017-15095	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2017-15095
CVE-2017-17485	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2017-17485
CVE-2017-7525	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2017-7525
CVE-2018-11307	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-11 307
CVE-2018-12022	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2018-12022
CVE-2018-12023	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-12 023
CVE-2018-14718	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14 718
CVE-2018-14719	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14 719
CVE-2018-14720	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14 720
CVE-2018-14721	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14 721
CVE-2018-19360	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-19360

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2018-19361	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-1936 1
CVE-2018-19362	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-1936 2
CVE-2018-5968	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-5968
CVE-2018-7489	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-7 489
CVE-2019-10172	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-10 172
CVE-2019-10202	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-10 202
CVE-2019-12086	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-12 2086
CVE-2019-14379	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14 379
CVE-2019-14439	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-14 439
CVE-2019-14540	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14 540
CVE-2019-14892	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14 892
CVE-2019-14893	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14 893
CVE-2019-16335	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16 335

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2019-16942	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16 942
CVE-2019-16943	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16 943
CVE-2019-17267	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17 267
CVE-2019-17531	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17 531
CVE-2019-20330	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20330
CVE-2020-10650	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10650
CVE-2020-10672	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10 672
CVE-2020-10673	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10 673
CVE-2020-10968	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10 968
CVE-2020-10969	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10 969
CVE-2020-11111	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-1111
CVE-2020-11112	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-1112
CVE-2020-11113	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-11-13

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-11619	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-11619
CVE-2020-11620	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-11620
CVE-2020-14060	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14060
CVE-2020-14061	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14061
CVE-2020-14062	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14062
CVE-2020-14195	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14 195
CVE-2020-24616	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-24 616
CVE-2020-24750	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-24750
CVE-2020-35490	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-35490
CVE-2020-35491	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-35491
CVE-2020-35728	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-35728
CVE-2020-36179	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36179
CVE-2020-36180	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36180

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-36181	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36181
CVE-2020-36182	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36182
CVE-2020-36183	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36183
CVE-2020-36184	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36184
CVE-2020-36185	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36185
CVE-2020-36186	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36186
CVE-2020-36187	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36187
CVE-2020-36188	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-3618
CVE-2020-36189	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36189
CVE-2020-36518	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36518
CVE-2020-8840	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-8840
CVE-2020-9546	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9546
CVE-2020-9547	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9547

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-9548	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9548
CVE-2021-20190	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-2019 0
CVE-2021-31684	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-3 684
CVE-2021-33813	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-33813
CVE-2022-3171	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2022-3171	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-3509	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3510
CVE-2022-3510	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3510
CVE-2022-42003	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2003
CVE-2022-42004	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-4 2004
CVE-2023-1370	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-1370

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2023-25194	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-25194
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-34610	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-34610
CVE-2023-3635	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3635
CVE-2023-39410	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-43642	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2023-44981	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2012-0881	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2012-08 81
CVE-2013-4002	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2013-4002
CVE-2015-1832	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2015-1832
CVE-2015-5237	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2015-5237
CVE-2018-11307	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-11 307

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2018-12022	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2018-12022
CVE-2018-12023	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-12 023
CVE-2018-14718	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14 718
CVE-2018-14719	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14 719
CVE-2018-14720	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14 720
CVE-2018-14721	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14 721
CVE-2018-19360	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-19360
CVE-2018-19361	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-1936 1
CVE-2018-19362	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-1936 2
CVE-2018-7489	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-7 489
CVE-2019-12086	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-12 2086
CVE-2019-14379	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14 379
CVE-2019-14439	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-14 439

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2019-14540	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14 540
CVE-2019-14892	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14 892
CVE-2019-14893	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14 893
CVE-2019-16335	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16 335
CVE-2019-16942	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16 942
CVE-2019-16943	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16 943
CVE-2019-17267	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17 267
CVE-2019-17531	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17 531
CVE-2019-20330	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20330
CVE-2020-10650	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10650
CVE-2020-10672	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10 672
CVE-2020-10673	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10 673
CVE-2020-10968	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10 968

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-10969	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10 969
CVE-2020-11111	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-1111
CVE-2020-11112	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-1112
CVE-2020-11113	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-11-13
CVE-2020-11619	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-11619
CVE-2020-11620	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-11620
CVE-2020-13936	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-13936
CVE-2020-13949	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2020-14060	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14060
CVE-2020-14061	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14061
CVE-2020-14062	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14062
CVE-2020-14195	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14 195
CVE-2020-24616	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-24 616

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-24750	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-24750
CVE-2020-25649	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-25 649
CVE-2020-35490	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-35490
CVE-2020-35491	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-35491
CVE-2020-35728	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-35728
CVE-2020-36179	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36179
CVE-2020-36180	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36180
CVE-2020-36181	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36181
CVE-2020-36182	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36182
CVE-2020-36183	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36183
CVE-2020-36184	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36184
CVE-2020-36185	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36185
CVE-2020-36186	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36186

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-36187	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36187
CVE-2020-36188	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-3618
CVE-2020-36189	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36189
CVE-2020-36518	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36518
CVE-2020-8840	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-8840
CVE-2020-9546	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9546
CVE-2020-9547	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9547
CVE-2020-9548	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9548
CVE-2021-20190	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-2019 0
CVE-2021-31684	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-3 684
CVE-2021-34538	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-34 538
CVE-2021-34538	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-34 538
CVE-2021-35515	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-35515

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2021-35516	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-35516
CVE-2021-35517	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-35517
CVE-2021-36090	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-36090
CVE-2022-25647	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-25 647
CVE-2022-3171	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3510
CVE-2022-36364	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3636 4
CVE-2022-37865	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-37865
CVE-2022-37866	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3786
CVE-2022-39135	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-39135
CVE-2022-40149	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40149
CVE-2022-40150	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40150

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2022-42003	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2003
CVE-2022-42003	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2003
CVE-2022-42004	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2004
CVE-2022-42004	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2004
CVE-2022-45685	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-45 685
CVE-2022-45693	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-45693
CVE-2022-46337	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-46337
CVE-2022-46751	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-46751
CVE-2023-1370	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2023-1436	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2023-34453	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3445 3
CVE-2023-34454	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3445
CVE-2023-34455	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3445
CVE-2023-39410	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-43642	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2022-25647	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-25 647
CVE-2022-42003	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2003
CVE-2022-42889	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-4289
CVE-2022-46751	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-46751
CVE-2023-1370	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2023-20883	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2083
CVE-2023-25194	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-25194
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-34453	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3445 3
CVE-2023-34454	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3445
CVE-2023-34455	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3445
CVE-2023-39410	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-39410	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-39410	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-43642	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2023-5072	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-5072
CVE-2018-8768	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-8 768
CVE-2021-32797	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-32797
CVE-2021-32798	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-327 98
CVE-2022-24758	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-24758

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2023-24580	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-24580
CVE-2023-36053	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-36053
CVE-2023-41164	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-41164
CVE-2023-43665	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-43665
CVE-2023-46695	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-46695
CVE-2024-24680	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-24680
CVE-2016-5018	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2016-5018
CVE-2016-6796	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2016-6796
CVE-2019-0205	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-0205
CVE-2020-13949	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-43642	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2012-0881	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2012-08 81

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2013-4002	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2013-4002
CVE-2019-17566	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-17 56
CVE-2019-17566	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-17 56
CVE-2020-11988	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-1 1988
CVE-2021-33813	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-33813
CVE-2022-25647	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-25 647
CVE-2022-34169	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-34169
CVE-2022-40146	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40146
CVE-2022-41704	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-41704
CVE-2022-42890	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-428 90
CVE-2022-44729	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-44729
CVE-2022-44729	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-44729
CVE-2022-44729	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-44729

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2022-46337	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-46337
CVE-2023-4759	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-4759
CVE-2015-2156	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2015-2156
CVE-2017-18640	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2017-18 640
CVE-2017-18640	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2017-18 640
CVE-2019-16869	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-16 869
CVE-2019-20444	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20 44
CVE-2019-20445	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20 445
CVE-2021-37136	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-37136
CVE-2021-37137	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-37137
CVE-2022-1471	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-1471
CVE-2022-1471	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-1471
CVE-2022-25857	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-25 857

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2022-25857	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-25 857
CVE-2022-3171	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3510
CVE-2022-42003	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2003
CVE-2022-42004	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2004
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-36478	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-36478
CVE-2023-44981	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2023-44981	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2015-2080	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2015-2080
CVE-2015-4852	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2015-4852

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2015-6420	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2015-6420
CVE-2015-7501	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2015-7501
CVE-2017-1000034	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2017-1000034
CVE-2017-12612	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2017-12612
CVE-2017-15288	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2017-1528
CVE-2017-5637	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2017-5637
CVE-2017-7656	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2017-7656
CVE-2017-7657	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2017-7 657
CVE-2017-7658	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2017-7658
CVE-2017-7687	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2017-7 687
CVE-2017-9735	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2017-9735
CVE-2017-9790	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2017-9790
CVE-2018-11793	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-17 93

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2018-11804	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-18 04
CVE-2018-17190	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-17 190
CVE-2018-8012	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-8012
CVE-2019-0204	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-0204
CVE-2019-10099	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-1099
CVE-2020-17533	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-1753
CVE-2020-9480	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9480
CVE-2021-28165	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-28 165
CVE-2021-35515	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-35515
CVE-2021-35515	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-35515
CVE-2021-35516	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-35516
CVE-2021-35516	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-35516
CVE-2021-35517	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-35517

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2021-35517	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-35517
CVE-2021-36090	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-36090
CVE-2021-36090	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-36090
CVE-2022-25647	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-25 647
CVE-2022-3171	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3510
CVE-2022-37866	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3786
CVE-2022-46751	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-46751
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-34453	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3445 3
CVE-2023-34454	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3445
CVE-2023-34455	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3445

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2023-43642	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2023-44981	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2014-0114	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2014-01 14
CVE-2016-1000338	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2016-1000338
CVE-2016-1000340	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2016-1000340
CVE-2016-1000342	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2016-1000342
CVE-2016-1000343	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2016-1000343
CVE-2016-1000344	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2016-1000344
CVE-2016-1000352	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2016-1000352
CVE-2017-1000487	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2017-1000487
CVE-2017-12629	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2017-12 629
CVE-2018-10936	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-10 936
CVE-2018-3258	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-3258

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2018-3827	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-3 827
CVE-2018-3831	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-38 31
CVE-2019-0205	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-0205
CVE-2019-10086	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-10 086
CVE-2019-10086	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-10 086
CVE-2019-7611	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-7611
CVE-2020-13692	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-13692
CVE-2020-13949	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2020-28491	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-28491
CVE-2020-7692	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-7692
CVE-2021-22573	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-22573
CVE-2021-22573	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-22573
CVE-2021-25646	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-25 646

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2021-26291	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-26291
CVE-2021-26291	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-26291
CVE-2021-26919	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-26919
CVE-2021-35515	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-35515
CVE-2021-35516	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-35516
CVE-2021-35517	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-35517
CVE-2021-36090	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-36090
CVE-2022-1471	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-1471
CVE-2022-25647	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-25 647
CVE-2022-25857	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-25 857
CVE-2022-31197	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-31 197
CVE-2022-3171	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2022-3171	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3171

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2022-33980	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-33980
CVE-2022-3509	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-3509	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3510
CVE-2022-3510	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3510
CVE-2022-42003	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2003
CVE-2022-42004	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2004
CVE-2022-4244	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-424
CVE-2022-45688	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-4568
CVE-2023-1370	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2023-1428	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-1428
CVE-2023-25194	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-25194
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-31418	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-314 18
CVE-2023-32731	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-32731
CVE-2023-34453	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3445 3
CVE-2023-34454	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3445
CVE-2023-34455	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3445
CVE-2023-3635	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3635
CVE-2023-38647	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-38647
CVE-2023-39410	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-43642	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2023-44981	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2023-46120	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-46120

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2023-5072	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-5072
CVE-2023-6378	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-6378
CVE-2023-6481	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-6481
CVE-2024-1597	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-1597
CVE-2024-1597	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-1597
CVE-2021-35958	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-35958
CVE-2019-0205	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-0205
CVE-2020-13949	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2021-34538	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-34 538
CVE-2022-45688	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-4568
CVE-2022-46337	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-46337
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-3635	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3635

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2023-5072	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-5072
CVE-2024-21634	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2024-21634
CVE-2021-35515	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-35515
CVE-2021-35516	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-35516
CVE-2021-35517	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-35517
CVE-2021-36090	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-36090
CVE-2022-40149	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40149
CVE-2022-40150	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40150
CVE-2022-45685	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-45 685
CVE-2022-45693	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-45693
CVE-2023-1436	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2014-0114	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2014-01 14
CVE-2015-2156	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2015-2156

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2018-10054	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-10 054
CVE-2019-0205	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-0205
CVE-2019-10086	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-10 086
CVE-2019-10086	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-10 086
CVE-2019-16869	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-16 869
CVE-2019-17195	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17 195
CVE-2019-20444	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20 44
CVE-2019-20445	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20 445
CVE-2020-13949	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2020-9492	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-9492
CVE-2021-37136	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-37136
CVE-2021-37137	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-37137
CVE-2021-37404	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-37404

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2022-25168	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-25 168
CVE-2022-25647	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-25 647
CVE-2022-25647	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-25 647
CVE-2022-26612	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-26612
CVE-2022-3171	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-3510
CVE-2022-40152	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40152
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-34453	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3445 3
CVE-2023-34454	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3445
CVE-2023-34455	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3445

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2023-3635	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3635
CVE-2023-3635	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-3635
CVE-2023-38647	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-38647
CVE-2023-43642	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2023-6378	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-6378
CVE-2023-6481	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-6481
CVE-2024-1597	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-1597
CVE-2017-15095	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2017-15095
CVE-2017-17485	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2017-17485
CVE-2018-11307	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-11 307
CVE-2018-12022	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2018-12022
CVE-2018-12023	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-12 023
CVE-2018-14718	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14 718

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2018-14719	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14 719
CVE-2018-14720	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14 720
CVE-2018-14721	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14 721
CVE-2018-19360	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-19360
CVE-2018-19361	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-1936 1
CVE-2018-19362	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-1936 2
CVE-2018-5968	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2018-5968
CVE-2018-7489	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-7 489
CVE-2019-12086	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-12 2086
CVE-2019-12086	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-12 2086
CVE-2019-14379	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14 379
CVE-2019-14379	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14 379
CVE-2019-14439	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-14 439

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2019-14439	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-14 439
CVE-2019-14540	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14 540
CVE-2019-14540	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14 540
CVE-2019-14892	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14 892
CVE-2019-14892	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14 892
CVE-2019-14893	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14 893
CVE-2019-16335	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16 335
CVE-2019-16335	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16 335
CVE-2019-16942	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16 942
CVE-2019-16942	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16 942
CVE-2019-16943	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16 943
CVE-2019-16943	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16 943
CVE-2019-17195	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17 195

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2019-17267	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17 267
CVE-2019-17267	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17 267
CVE-2019-17531	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17 531
CVE-2019-17531	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17 531
CVE-2019-20330	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20330
CVE-2019-20330	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20330
CVE-2020-10650	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10650
CVE-2020-10650	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10650
CVE-2020-10672	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10 672
CVE-2020-10672	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10 672
CVE-2020-10673	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10 673
CVE-2020-10673	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10 673
CVE-2020-10968	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10 968

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-10968	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10 968
CVE-2020-10969	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10 969
CVE-2020-10969	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-10 969
CVE-2020-11111	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-1111
CVE-2020-11111	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-1111
CVE-2020-11112	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-1112
CVE-2020-11113	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-11-13
CVE-2020-11619	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-11619
CVE-2020-11620	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-11620
CVE-2020-11620	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-11620
CVE-2020-14060	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14060
CVE-2020-14060	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14060
CVE-2020-14061	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14061

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-14061	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14061
CVE-2020-14062	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14062
CVE-2020-14062	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14062
CVE-2020-14195	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14 195
CVE-2020-14195	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-14 195
CVE-2020-24616	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-24 616
CVE-2020-24616	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-24 616
CVE-2020-24750	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-24750
CVE-2020-24750	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-24750
CVE-2020-25649	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-25 649
CVE-2020-25649	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-25 649
CVE-2020-25649	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-25 649
CVE-2020-35211	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-3521

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-35490	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-35490
CVE-2020-35490	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-35490
CVE-2020-35490	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-35490
CVE-2020-35491	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-35491
CVE-2020-35491	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-35491
CVE-2020-35491	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-35491
CVE-2020-35728	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-35728
CVE-2020-35728	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-35728
CVE-2020-35728	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-35728
CVE-2020-36179	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36179
CVE-2020-36179	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36179
CVE-2020-36179	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36179
CVE-2020-36180	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36180

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-36180	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36180
CVE-2020-36180	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36180
CVE-2020-36181	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36181
CVE-2020-36181	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36181
CVE-2020-36181	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36181
CVE-2020-36182	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36182
CVE-2020-36182	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36182
CVE-2020-36182	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36182
CVE-2020-36183	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36183
CVE-2020-36183	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36183
CVE-2020-36183	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36183
CVE-2020-36184	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36184
CVE-2020-36184	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36184

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-36184	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36184
CVE-2020-36185	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36185
CVE-2020-36185	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36185
CVE-2020-36185	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36185
CVE-2020-36186	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36186
CVE-2020-36186	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36186
CVE-2020-36186	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36186
CVE-2020-36187	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36187
CVE-2020-36187	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36187
CVE-2020-36187	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36187
CVE-2020-36188	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-3618
CVE-2020-36188	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-3618
CVE-2020-36188	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-3618

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-36189	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36189
CVE-2020-36189	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36189
CVE-2020-36189	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36189
CVE-2020-36518	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36518
CVE-2020-36518	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36518
CVE-2020-36518	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36518
CVE-2020-36518	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2020-36518
CVE-2020-7692	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-7692
CVE-2020-8840	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-8840
CVE-2020-8840	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-8840
CVE-2020-9546	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9546
CVE-2020-9546	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9546
CVE-2020-9547	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9547

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2020-9547	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9547
CVE-2020-9548	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9548
CVE-2020-9548	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9548
CVE-2021-20190	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-2019 0
CVE-2021-20190	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-2019 0
CVE-2021-20190	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-2019 0
CVE-2021-22573	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-22573
CVE-2021-33036	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-330 36
CVE-2021-35515	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-35515
CVE-2021-35515	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-35515
CVE-2021-35516	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-35516
CVE-2021-35516	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-35516
CVE-2021-37714	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-377 14

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2021-37714	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-377 14
CVE-2021-37714	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-377 14
CVE-2021-37714	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-377 14
CVE-2021-46877	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-4687
CVE-2022-25647	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-25 647
CVE-2022-32532	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-32532
CVE-2022-40149	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40149
CVE-2022-40150	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40150
CVE-2022-40664	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-4064
CVE-2022-42003	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2003
CVE-2022-42003	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2003
CVE-2022-42003	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2003
CVE-2022-42003	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2003

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2022-42004	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2004
CVE-2022-42004	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2004
CVE-2022-42004	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2004
CVE-2022-42004	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE- 2022-4 2004
CVE-2022-45685	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-45 685
CVE-2022-45688	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-4568
CVE-2022-45693	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-45693
CVE-2023-1436	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2022-40149	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40149
CVE-2022-40150	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-40150
CVE-2022-45685	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-45 685

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2022-45693	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2022-45693
CVE-2023-1436	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2023-6378	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-6378
CVE-2023-6378	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-6378
CVE-2023-6481	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-6481
CVE-2023-6481	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2023-6481

Vulnerabilidades e exposições comuns corrigidas em mecanismos centrais da versão 7.2.0

A tabela a seguir lista todas as correções CVEs no Amazon EMR 7.2.0.

CVE ID	Gravidade	URL de detalhes da CVE
CVE-2019-12402	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2019-12 402
CVE-2021-41561	HIGH (ALTO)	https://nvd.nist. gov/vuln/detail/CVE-2021-41561

Detalhes dos componentes da versão 7.2.0

Consulte a tabela a seguir para obter mais informações sobre os pacotes Extras no Amazon EMR 7.2.0.

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
AmazonCloudWatchAgent	Serviços	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.4-1.amzn2023
AmazonCloudWatchAgent	Serviços	ranger-kms	Extras	2.0.0-1.amzn2023
AmazonCloudWatchAgent	Serviços	s3-dist-cp	Extras	2.32.0-1.amzn2023
AmazonCloudWatchAgent	Serviços	zookeeper	Extras	3.9.1.amzn.1-1.amzn2023
AmazonCloudWatchAgent	Serviços	zookeeper-server	Extras	3.9.1.amzn.1-1.amzn2023
Flink	Serviços	hadoop-client	Extras	3.3.6.amzn.4-1.amzn2023
Flink	Serviços	hadoop-hdfs	Extras	3.3.6.amzn.4-1.amzn2023
Flink	Serviços	hadoop-hdfs-datanode	Extras	3.3.6.amzn.4-1.amzn2023
Flink	Serviços	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.4-1.amzn2023
Flink	Serviços	hadoop-hdfs-namenode	Extras	3.3.6.amzn.4-1.amzn2023
Flink	Serviços	hadoop-httpfs	Extras	3.3.6.amzn.4-1.amzn2023
Flink	Serviços	hadoop-kms	Extras	3.3.6.amzn.4-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Flink	Serviços	hadoop-mapreduce	Extras	3.3.6.amzn.4-1.amzn2023
Flink	Serviços	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.4-1.amzn2023
Flink	Serviços	hadoop-yarn	Extras	3.3.6.amzn.4-1.amzn2023
Flink	Serviços	hadoop-yarn-nodemanager	Extras	3.3.6.amzn.4-1.amzn2023
Flink	Serviços	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.4-1.amzn2023
Flink	Serviços	hadoop-yarn-resourcemanager	Extras	3.3.6.amzn.4-1.amzn2023
Flink	Serviços	hadoop-yarn-timelineserver	Extras	3.3.6.amzn.4-1.amzn2023
Flink	Serviços	hudi	Extras	0.14.1.amzn.1-1.amzn2023
Flink	Serviços	ranger-kms	Extras	2.0.0-1.amzn2023
Flink	Serviços	hadoop	Extras	3.3.6.amzn.4-1.amzn2023
HBase	Serviços	s3-dist-cp	Extras	2.32.0-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
HCatalog	Serviços	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.4-1.amzn2023
HCatalog	Serviços	mariadb-connector-java	Extras	2.7.2-1
Hive	Serviços	hadoop-httpfs	Extras	3.3.6.amzn.4-1.amzn2023
Hive	Serviços	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.4-1.amzn2023
Hive	Serviços	hive-hbase	Extras	3.1.3.amzn.11-1.amzn2023
Hive	Serviços	mariadb-connector-java	Extras	2.7.2-1
Hive	Serviços	s3-dist-cp	Extras	2.32.0-1.amzn2023
Phoenix	Serviços	s3-dist-cp	Extras	2.32.0-1.amzn2023
Spark	Serviços	spark-datanucleus	Extras	3.5.1.amzn.0-1.amzn2023
Tez	Serviços	hadoop-mapreduce	Extras	3.3.6.amzn.4-1.amzn2023
Tez	Serviços	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.4-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Tez	Serviços	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Serviços	bigtop-utils	Extras	1.2.0-1.amzn2023
Trino	Serviços	hadoop-client	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Serviços	hadoop-hdfs	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Serviços	hadoop-hdfs-datanode	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Serviços	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Serviços	hadoop-hdfs-namenode	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Serviços	hadoop-kms	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Serviços	hadoop-mapreduce	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Serviços	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Serviços	hadoop-yarn	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Serviços	hadoop-yarn-nodemanager	Extras	3.3.6.amzn.4-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Trino	Serviços	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Serviços	hadoop-yarn-resourcemanager	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Serviços	hadoop-yarn-timelineserver	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Serviços	hive	Extras	3.1.3.amzn.11-1.amzn2023
Trino	Serviços	hive-hcatalog-server	Extras	3.1.3.amzn.11-1.amzn2023
Trino	Serviços	mariadb-connector-java	Extras	2.7.2-1
Trino	Serviços	ranger-kms	Extras	2.0.0-1.amzn2023
Trino	Serviços	zookeeper	Extras	3.9.1.amzn.1-1.amzn2023
Trino	Serviços	zookeeper-server	Extras	3.9.1.amzn.1-1.amzn2023
Trino	Serviços	bigtop-groovy	Extras	2.5.4-1.amzn2023
Trino	Serviços	bigtop-jsvc	Extras	1.2.4-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Trino	Serviços	hadoop	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Serviços	hadoop-lzo	Extras	0.4.19-1.amzn2023
Trino	Serviços	hive-hcatalog	Extras	3.1.3.amzn.11-1.amzn2023
Trino	Serviços	hive-jdbc	Extras	3.1.3.amzn.11-1.amzn2023
Zookeeper	Serviços	emrfs	Extras	2.63.0-1.amzn2023
Zookeeper	Serviços	hadoop-client	Extras	3.3.6.amzn.4-1.amzn2023
Zookeeper	Serviços	hadoop-hdfs	Extras	3.3.6.amzn.4-1.amzn2023
Zookeeper	Serviços	hadoop-hdfs-datanode	Extras	3.3.6.amzn.4-1.amzn2023
Zookeeper	Serviços	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.4-1.amzn2023
Zookeeper	Serviços	hadoop-hdfs-namenode	Extras	3.3.6.amzn.4-1.amzn2023
Zookeeper	Serviços	hadoop-httpfs	Extras	3.3.6.amzn.4-1.amzn2023

Aplicação	Tipo de aplicação	Pacote	Tipo de pacote	Versão do pacote
Zookeeper	Serviços	hadoop-kms	Extras	3.3.6.amzn.4-1.amzn2023
Zookeeper	Serviços	hadoop-mapreduce	Extras	3.3.6.amzn.4-1.amzn2023
Zookeeper	Serviços	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.4-1.amzn2023
Zookeeper	Serviços	hadoop-yarn	Extras	3.3.6.amzn.4-1.amzn2023
Zookeeper	Serviços	hadoop-yarn-nodemanager	Extras	3.3.6.amzn.4-1.amzn2023
Zookeeper	Serviços	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.4-1.amzn2023
Zookeeper	Serviços	hadoop-yarn-resourcemanager	Extras	3.3.6.amzn.4-1.amzn2023
Zookeeper	Serviços	hadoop-yarn-timelineserver	Extras	3.3.6.amzn.4-1.amzn2023
Zookeeper	Serviços	ranger-kms	Extras	2.0.0-1.amzn2023

Notas da versão 7.2.0

As notas da versão a seguir incluem informações sobre a versão 7.2.0 do Amazon EMR. As alterações são referentes à versão 7.2.0.

Novos recursos

- Atualizações da aplicação: Amazon EMR 7.2.0 application upgrades include Iceberg 1.5.0-amzn-0 and Delta 3.1.0.
- O Amazon EMR adiciona suporte para que você possa usar outros aplicativos HBase, como Flink e Hive, com a classe de armazenamento Amazon S3 Express One Zone.
- Essa versão adiciona a capacidade de ler objetos restaurados, para que você possa ler objetos Glacier de um local do S3 com o protocolo `S3A`. Esse recurso funciona com Spark, Flink e Hive.
- Rótulos de nós com ajuste de escala gerenciado: use rótulos de nós com instâncias de acordo com o tipo de mercado ou o tipo de nó para melhorar o ajuste de escala automático do Amazon EMR. Para obter mais informações, consulte Using managed scaling in Amazon EMR.

Problemas conhecidos

- Não há suporte para o Python 3.11 no EMR Studio.

Alterações, melhorias e problemas resolvidos

- Essa versão corrige um problema de deadlock que pode ocorrer durante as operações internas de limpeza de etapas. Essa operação gerencia o ciclo de vida das etapas à medida que elas são concluídas no cluster do EMR. Esse problema afeta operações críticas do Amazon EMR, como operação por etapas e ajuste de escala.
- Esta versão resolve um problema em que clusters personalizados com determinados arquivos de log preexistentes podem fazer com AMIs que o daemon de gerenciamento de log do Amazon EMR falhe.
- O Amazon EMR 7.2.0 atualiza o daemon do Amazon EMR responsável pelas atividades de gerenciamento e monitoramento de clusters do SDK v1 para v2. AWS

- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250527.1	6.1.134-152.225.amzn2023	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein), China (Pequim), Ásia-Pacífico (Malásia), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Leste dos EUA (Norte da Virgínia), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Malásia), Europa (Londres), Ásia-Pacífico (Melbourne), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (AWS GovCloud AWS GovCloud Oregon), México (Central), Oeste dos EUA (Norte da

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Califórnia), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv), Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250512.(	6.1.134-152.225.am zn2023	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Hyderabad), Ásia-Pacífico (Jacarta)), Ásia-Pacífico (Malásia), Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Oriente Médio (Oriente Médio) (Oriente Médio (Oriente Médio)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Emirados Árabes Unidos), Canadá (Central) , Oeste do Canadá (Calgary), (Leste dos EUA), AWS GovCloud (AWS GovCloud Oeste dos EUA), China (Pequim), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250428.1	6.1.134-150.224.amzn2023	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Jacarta) Arta), Ásia-Pacífico (Malásia), Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Malásia) , Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (EAU), Canadá (Central) , Oeste do Canadá

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Calgary), (Leste dos EUA), AWS GovCloud (Oeste dos EUA), China (AWS GovCloud Pequim), China (Ningxia), Israel (Tel Aviv), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250414.(	6.1.132-147.221.am zn2023	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250331.(	6.1.131-143.221.am zn2023	18 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central), Ásia-Pacífico (Tóquio)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250303.(	6.1.129-138.220.am zn2023	27 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250218.1	6.1.128-136.201.amzn2023	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderabad), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250211.(	6.1.127-135.201.am zn2023	26 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250123.4	6.1.124-134.200.amzn2023	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderabad), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250115.(	6.1.119-129.201.am zn2023	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 241121.(	6.1.115-126.197.am zn2023	12 de dezembro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 241031.(	6.1.112-124.190.am zn2023	15 de novembro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.5.2 241001.1	6.1.109-118.189.am zn2023	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.5.2 240819.(	6.1.102-111.182.am zn2023	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.5.2 240730.(	6.1.97-104.177.amzn2023	2 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.5.2 240722.(	6.1.97-104.177.amzn2023	24 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.5.2 240708.(	6.1.96-102.177.amzn2023	8 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), Oeste

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			do Canadá (Calgary) , (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China AWS GovCloud (Pequim), China (Ningxia)

Versões padrão do Java 7.2.0

O Amazon EMR 7.2 e versões superiores vêm com o Amazon Corretto 17 (criado no OpenJDK) por padrão para aplicações compatíveis com o Corretto 17 (JDK 17), com exceção do Apache Livy.

A tabela a seguir mostra as versões padrão do Java para aplicações no Amazon EMR 7.2.0. Se você quiser alterar a JVM padrão no cluster, siga as instruções em Configurar aplicações para usar uma Máquina Virtual Java específica para cada aplicação executada no cluster. Você só pode usar uma versão de runtime do Java para um cluster. O Amazon EMR não é compatível com a execução de diferentes nós ou aplicações em diferentes versões de runtime no mesmo cluster.

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
Delta	17, 11, 8
Flink	17, 11, 8
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez	17, 11, 8
Trino	21, 17
Zeppelin	8
Pig	8
Zookeeper	8

Versões de componente 7.2.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *CommunityVersion*-amzn-*EmrVersion*. O *EmrVersion* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado myapp-component com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como 2.2-amzn-2.

Componente	Versão	Descrição
adot-java-agent	1.31.0	Um Java Agent que coleta métricas de daemons de aplicações.
delta	3.1.0	O Delta Lake é um formato de tabela aberto para grandes conjuntos de dados analíticos
emr-amazon-cloudwatch-agent	1.300032.2-amzn-0	Um aplicativo que coleta métricas internas em nível de sistema e métricas de aplicativos personalizadas de instâncias da Amazon. EC2
emr-ddb	5.4.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.11.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.15.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.16.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway

Componente	Versão	Descrição
emr-s3-dist-cp	2.32.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	2.11.0	EMR S3Select Connector
emr-wal-cli	1.2.2	Cli usada para lista/exclusão de emrwal.
emrfs	2.63.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.18.1-amzn-1	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.18.1-amzn-1	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
hadoop-client	3.3.6-amzn-4	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	3.3.6-amzn-4	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.3.6-amzn-4	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.3.6-amzn-4	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.

Componente	Versão	Descrição
hadoop-hdfs-journalnode	3.3.6-amzn-4	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.3.6-amzn-4	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.3.6-amzn-4	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	3.3.6-amzn-4	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.3.6-amzn-4	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	3.3.6-amzn-4	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	3.3.6-amzn-4	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	2.4.17-amzn-6	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.

Componente	Versão	Descrição
hbase-region-server	2.4.17-amzn-6	Serviço para atender a uma ou mais HBase regiões.
hbase-client	2.4.17-amzn-6	HBase cliente de linha de comando.
hbase-rest-server	2.4.17-amzn-6	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.4.17-amzn-6	Serviço que fornece um endpoint Thrift para. HBase
hbase-operator-tools	2.4.17-amzn-6	Ferramenta de reparo para HBase clusters Apache.
hcatalog-client	3.1.3-amzn-11	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.3-amzn-11	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	3.1.3-amzn-11	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.3-amzn-11	O cliente da linha de comando do Hive.
hive-hbase	3.1.3-amzn-11	Cliente do Hive-hbase.

Componente	Versão	Descrição
hive-metastore-server	3.1.3-amzn-11	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.3-amzn-11	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.14.1-amzn-1	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.14.1-amzn-1	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-trino	0.14.1-amzn-1	Biblioteca de pacotes para execução do Presto com o Hudi.
hudi-spark	0.14.1-amzn-1	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.11.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
iceberg	1.5.0-amzn-0	Apache Iceberg é um formato de tabela aberta para conjuntos de dados analíticos imensos
jupyterhub	1.5.0	Servidor de vários usuários para blocos de anotações Jupyter

Componente	Versão	Descrição
livy-server	0.8.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mxnet	1.9.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.
nvidia-cuda	11.8.0	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.7.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	5.1.3	As bibliotecas do phoenix para servidor e cliente
phoenix-connectors	5.1.3	Conectores do Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica

Componente	Versão	Descrição
presto-coordinator	0.285-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.285-amzn-0	O serviço que executa partes de uma consulta.
presto-client	0.285-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
trino-coordinator	436-amzn-0	Serviço que aceita consultas e gerencia a execução de consultas entre trino-workers.
trino-worker	436-amzn-0	O serviço que executa partes de uma consulta.
trino-client	436-amzn-0	Cliente de linha de comando do Trino instalado em mestres de espera de um cluster de alta disponibilidade (HA) em que o servidor do Trino não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.3.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger

Componente	Versão	Descrição
spark-client	3.5.1-amzn-0	Os clientes da linha de comando do Spark.
spark-history-server	3.5.1-amzn-0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.5.1-amzn-0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.5.1-amzn-0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
spark-rapids	24.02.0-amzn-0	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.11.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.10.2-amzn-9	O aplicativo e as bibliotecas do YARN para Tez.
tez-on-worker	0.10.2-amzn-9	A aplicação e as bibliotecas do YARN nós de processam ento.

Componente	Versão	Descrição
zeppelin-server	0.10.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.9.1-amzn-1	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.9.1-amzn-1	ZooKeeper cliente de linha de comando.

Classificações de configuração 7.2.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-7.2.0

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.

Classificações	Descrição	Ações de reconfiguração
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.

Classificações	Descrição	Ações de reconfiguração
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.

Classificações	Descrição	Ações de reconfiguração
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classificações	Descrição	Ações de reconfiguração
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.

Classificações	Descrição	Ações de reconfiguração
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Alterar a página inicial do KMS do Hadoop	Not available.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.	Not available.
iceberg-defaults	Alterar os valores no arquivo iceberg-defaults.conf do Iceberg.	Not available.
delta-defaults	Alterar os valores no arquivo delta-defaults.conf do Delta.	Not available.

Classificações	Descrição	Ações de reconfiguração
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j2	Alterar as configurações de log4j2.properties do Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.

Classificações	Descrição	Ações de reconfiguração
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-lakeformation	Alterar os valores no arquivo lakeformation.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
trino-log	Alterar os valores no arquivo log.properties do Trino.	Restarts Trino-Server (for Trino)
trino-config	Alterar os valores no arquivo config.properties do Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Alterar os valores no arquivo password-authenticator.properties do Trino.	Restarts Trino-Server (for Trino)
trino-env	Alterar os valores no arquivo trino-env.sh do Trino.	Restarts Trino-Server (for Trino)

Classificações	Descrição	Ações de reconfiguração
trino-node	Alterar os valores no arquivo node.properties do Trino.	Not available.
trino-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Trino.	Not available.
trino-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Trino.	Not available.
trino-connector-delta	Alterar os valores no arquivo delta.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Alterar os valores no arquivo hive.properties do Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Alterar os valores no arquivo exchange-manager.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Alterar os valores no arquivo iceberg.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Alterar os valores no arquivo hudi.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Alterar os valores no arquivo jmx.properties do Trino.	Not available.
trino-connector-kafka	Alterar os valores no arquivo kafka.properties do Trino.	Not available.
trino-connector-localfile	Alterar os valores no arquivo localfile.properties do Trino.	Not available.
trino-connector-memory	Alterar os valores no arquivo memory.properties do Trino.	Not available.

Classificações	Descrição	Ações de reconfiguração
trino-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Trino.	Not available.
trino-connector-mysql	Alterar os valores no arquivo mysql.properties do Trino.	Not available.
trino-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Trino.	Not available.
trino-connector-raptor	Alterar os valores no arquivo raptor.properties do Trino.	Not available.
trino-connector-redis	Alterar os valores no arquivo redis.properties do Trino.	Not available.
trino-connector-redshift	Alterar os valores no arquivo redshift.properties do Trino.	Not available.
trino-connector-tpch	Alterar os valores no arquivo tpch.properties do Trino.	Not available.
trino-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Trino.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Alterar os valores no arquivo kms-logback.xml do Ranger KMS.	Not available.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
spark	As configurações administr adas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j2	Alterar os valores no arquivo log4j2.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie and HiveServe r2.

Classificações	Descrição	Ações de reconfiguração
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.
zookeeper-logback	Altere os valores no ZooKeeper arquivo logback.xml do.	Restarts Zookeeper server.
emr-metrics	Altere as configurações de métricas do EMR desse nó.	Restarts the CloudWatchAgent service.

Log de alterações 7.2.0

Log de alterações da versão 7.2.0 e das notas da versão

Data	Event	Descrição
2024-07-25	Publicação de documentos	Notas da versão 7.2.0 do Amazon EMR publicadas pela primeira vez
2024-07-16	Lançamento inicial	A versão 7.2.0 do Amazon EMR foi implantada pela primeira vez em regiões comerciais iniciais

Versão 7.1.0 do Amazon EMR

Versões 7.1.0 da aplicação

Essa versão inclui os seguintes aplicativos: AmazonCloudWatchAgent, Delta, Flink, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, SqoopTensorFlow, Tez, TrinoZeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
AWS SDK for Java	2.25.70, 1.12.747	2.23.18, 1.12.705	2.23.18, 1.12.656	2.20.160-amzn-0, 1.12.569
Python	3,9, 3,11	3,9, 3,11	3,9, 3,11	3.9
Scala	2.12.18	2.12.17	2.12.17	2.12.17
AmazonCloudWatchAgent	1.300032.2	1.300032.2	1.300032.2	1.300031.1
Delta	3.2.0-amzn-0	3.1.0	3.0.0	3.0.0
Flink	1.18.1	1.18.1	1.18.1	1.18.0
Ganglia	-	-	-	-
HBase	2.4.17	2.4.17	2.4.17	2.4.17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.6	3.3.6	3.3.6	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-amzn-0	0.14.1-amzn-1	0.14.1-amzn-0	0.14.0-amzn-1
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.5.2-amzn-0	1.5.0-amzn-0	1.4.3-amzn-0	q1.4.2-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.7.1

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
MXNet	-	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,285	0,285	0,284	0,283
Spark	3.5.1	3.5.1	3.5.0	3.5.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.16.1	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	442	436	435	426
Zeppelin	0.11.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.9.1	3.9.1	3.9.1	3.5.10

Notas da versão 7.1.0

As notas da versão a seguir incluem informações sobre a versão 7.1.0 do Amazon EMR. As alterações são referentes à versão 7.1.0.

Novos recursos

- Atualizações da aplicação: Amazon EMR 7.1.0 application upgrades include Livy 0.8.0, Trino 435, and ZooKeeper 3.9.1.
- Substituição de nós não íntegros: com o Amazon EMR 7.1.0 e versões posteriores, a substituição de nós não íntegros é habilitada por padrão, portanto, o Amazon EMR substituirá os nós não

íntegros normalmente. Para evitar afetar seus fluxos de trabalho existentes nas versões 7.0.0 e inferiores do Amazon EMR, a substituição de nós não íntegros será desabilitada se você habilitar a proteção contra encerramento no cluster.

- CloudWatch Agente — Configure o CloudWatch agente para usar métricas adicionais do sistema, adicionar métricas de aplicativos e alterar o destino das métricas com a API de configuração do Amazon EMR.

Problemas conhecidos

- O aplicativo JupyterEnterpriseGateway (JEG) não oferece suporte ao Python 3.11 nesta versão.

Alterações, melhorias e problemas resolvidos

- Enquanto o Amazon EMR 7.1.0 oferece suporte ao Python 3.9 por padrão, o Livy 0.8.0 e o Spark no Amazon EMR 7.1.0 oferecem suporte ao Python 3.11.
- Esta versão corrige o problema da necessidade de executar cada linha, uma por vez, ao usar PySpark com o Python versão 3.11.
- Atualização do Zeppelin — O Amazon EMR 7.1.0 inclui uma atualização do Zeppelin para o SDK for Java v2. AWS Essa atualização permite que um caderno S3 do Zeppelin aceite um fornecedor de materiais de criptografia personalizados. O AWS SDK for Java v2 remove EncryptionMaterialsProvider a interface. Ao fazer o upgrade para o Amazon EMR 7.1.0, você deve implementar a interface de chaveiro se quiser usar criptografia personalizada. Para ver um exemplo de como implementar a interface Keyring, consulte KmsKeyring.java.
- Ao fazer o upgrade para a versão 7.1.0 do Amazon EMR, altere seu provedor de chaves personalizadas para criptografia de disco local para gerar chaves usando o algoritmo AES de. AES/GCM/NoPadding Se você não atualizar o algoritmo, a criação do cluster poderá falhar com o erro `Local disk encryption failed on master instance (i-123456789) due to internal error.`. Para obter mais informações sobre como criar um provedor de chave personalizada, consulte Creating a custom key provider.
- O Amazon EMR 7.1.0 melhora a resiliência de um nó em condições de pouco espaço em disco, otimizando a lógica de truncamento de log para arquivos com identificadores de arquivo abertos.
- Essa versão aprimora a lógica de codificação e decodificação para minimizar o risco de corrupção de dados e falha de nó com os daemons do Amazon EMR que leem e gravam arquivos ao reiniciar um nó.

- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250527.1	6.1.134-152.225.am zn2023	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Malásia), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Malásia), Europa (Londres), Ásia-Pacífico (Melbourne), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (AWS GovCloud AWS GovCloud Oregon), México (Central) , Oeste dos EUA (Norte da

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Califórnia), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv), Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250512.(	6.1.134-152.225.am zn2023	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Hyderabad), Ásia-Pacífico (Jacarta)), Ásia-Pacífico (Malásia), Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Oriente Médio (Oriente Médio) (Oriente Médio (Oriente Médio)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Emirados Árabes Unidos), Canadá (Central) , Oeste do Canadá (Calgary), (Leste dos EUA), AWS GovCloud (AWS GovCloud Oeste dos EUA), China (Pequim), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250428.1	6.1.134-150.224.amzn2023	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Jacarta) Arta), Ásia-Pacífico (Malásia), Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Malásia) , Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (EAU), Canadá (Central) , Oeste do Canadá

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Calgary), (Leste dos EUA), AWS GovCloud (Oeste dos EUA), China (AWS GovCloud Pequim), China (Ningxia), Israel (Tel Aviv), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250414.(	6.1.132-147.221.am zn2023	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250331.(	6.1.131-143.221.am zn2023	18 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central), Ásia-Pacífico (Tóquio)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250303.(	6.1.129-138.220.am zn2023	27 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250218.1	6.1.128-136.201.amzn2023	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderabad), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250211.(	6.1.127-135.201.am zn2023	26 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250123.4	6.1.124-134.200.amzn2023	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderabad), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250115.(	6.1.119-129.201.am zn2023	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 241121.(	6.1.115-126.197.am zn2023	12 de dezembro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 241031.(	6.1.112-124.190.am zn2023	15 de novembro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.5.2 241001.1	6.1.109-118.189.am zn2023	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.5.2 240819.(	6.1.102-111.182.am zn2023	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.5.2 240730.(	6.1.97-104.177.amzn2023	2 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.5.2 240722.(	6.1.97-104.177.amzn2023	24 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.5.2 240708.(	6.1.96-102.177.amzn2023	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.3.2 240219.(	6.1.77-99.164.amzn2023	8 de maio de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), Oeste

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			do Canadá (Calgary) , (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China AWS GovCloud (Pequim), China (Ningxia)

Versões padrão do Java 7.1.0

O Amazon EMR 7.1 e versões superiores vêm com o Amazon Corretto 17 (criado no OpenJDK) por padrão para aplicações compatíveis com o Corretto 17 (JDK 17), com exceção do Apache Livy.

A tabela a seguir mostra as versões padrão do Java para aplicações no Amazon EMR 7.1.0. Se você quiser alterar a JVM padrão no cluster, siga as instruções em Configurar aplicações para usar uma Máquina Virtual Java específica para cada aplicação executada no cluster. Você só pode usar uma versão de runtime do Java para um cluster. O Amazon EMR não é compatível com a execução de diferentes nós ou aplicações em diferentes versões de runtime no mesmo cluster.

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
Delta	17, 11, 8
Flink	17, 11, 8
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez	17, 11, 8
Trino	17
Zeppelin	8
Pig	8
Zookeeper	8

Versões de componente 7.1.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *CommunityVersion*-amzn-*EmrVersion*. O *EmrVersion* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado myapp-component com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como 2.2-amzn-2.

Componente	Versão	Descrição
adot-java-agent	1.31.0	Um Java Agent que coleta métricas de daemons de aplicações.
delta	3.0.0	O Delta Lake é um formato de tabela aberto para grandes conjuntos de dados analíticos
emr-amazon-cloudwatch-agent	1.300032.2-amzn-0	Um aplicativo que coleta métricas internas em nível de sistema e métricas de aplicativos personalizadas de instâncias da Amazon. EC2
emr-ddb	5.3.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.10.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.14.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.7.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway

Componente	Versão	Descrição
emr-s3-dist-cp	2.31.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	2.10.0	EMR S3Select Connector
emr-wal-cli	1.2.2	Cli usada para lista/exclusão de emrwal.
emrfs	2.62.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.18.1-amzn-0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.18.1-amzn-0	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
hadoop-client	3.3.6-amzn-3	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	3.3.6-amzn-3	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.3.6-amzn-3	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.3.6-amzn-3	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.

Componente	Versão	Descrição
hadoop-hdfs-journalnode	3.3.6-amzn-3	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.3.6-amzn-3	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.3.6-amzn-3	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	3.3.6-amzn-3	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.3.6-amzn-3	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	3.3.6-amzn-3	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	3.3.6-amzn-3	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	2.4.17-amzn-5	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.

Componente	Versão	Descrição
hbase-region-server	2.4.17-amzn-5	Serviço para atender a uma ou mais HBase regiões.
hbase-client	2.4.17-amzn-5	HBase cliente de linha de comando.
hbase-rest-server	2.4.17-amzn-5	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.4.17-amzn-5	Serviço que fornece um endpoint Thrift para. HBase
hbase-operator-tools	2.4.17-amzn-5	Ferramenta de reparo para HBase clusters Apache.
hcatalog-client	3.1.3-amzn-10	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.3-amzn-10	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	3.1.3-amzn-10	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.3-amzn-10	O cliente da linha de comando do Hive.
hive-hbase	3.1.3-amzn-10	Cliente do Hive-hbase.

Componente	Versão	Descrição
hive-metastore-server	3.1.3-amzn-10	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.3-amzn-10	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.14.1-amzn-0	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.14.1-amzn-0	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-trino	0.14.1-amzn-0	Biblioteca de pacotes para execução do Presto com o Hudi.
hudi-spark	0.14.1-amzn-0	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.11.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
iceberg	1.4.3-amzn-0	Apache Iceberg é um formato de tabela aberta para conjuntos de dados analíticos imensos
jupyterhub	1.5.0	Servidor de vários usuários para blocos de anotações Jupyter

Componente	Versão	Descrição
livy-server	0.8.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mxnet	1.9.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.
nvidia-cuda	11.8.0	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.7.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	5.1.3	As bibliotecas do phoenix para servidor e cliente
phoenix-connectors	5.1.3	Conectores do Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica

Componente	Versão	Descrição
presto-coordinator	0.284-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.284-amzn-0	O serviço que executa partes de uma consulta.
presto-client	0.284-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
trino-coordinator	435-amzn-0	Serviço que aceita consultas e gerencia a execução de consultas entre trino-workers.
trino-worker	435-amzn-0	O serviço que executa partes de uma consulta.
trino-client	435-amzn-0	Cliente de linha de comando do Trino instalado em mestres de espera de um cluster de alta disponibilidade (HA) em que o servidor do Trino não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.3.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger

Componente	Versão	Descrição
spark-client	3.5.0-amzn-1	Os clientes da linha de comando do Spark.
spark-history-server	3.5.0-amzn-1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.5.0-amzn-1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.5.0-amzn-1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
spark-rapids	23.10.0-amzn-1	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.11.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.10.2-amzn-8	O aplicativo e as bibliotecas do YARN para Tez.
tez-on-worker	0.10.2-amzn-8	A aplicação e as bibliotecas do YARN nós de processam ento.

Componente	Versão	Descrição
zeppelin-server	0.10.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.9.1-amzn-0	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.9.1-amzn-0	ZooKeeper cliente de linha de comando.

Classificações de configuração 7.1.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-7.1.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.

Classificações	Descrição
docker-conf	Alterar as configurações relacionadas ao docker.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-session	Altere as configurações do Flink log4j-ses sion.properties para a sessão. Kubernetes/Yarn
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.

Classificações	Descrição
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-env	Alterar os valores no ambiente do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.

Classificações	Descrição
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.proper ties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.x ml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-java-home	Alterar a página inicial do KMS do Hadoop
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properti es do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.

Classificações	Descrição
hudi-env	Altere os valores no ambiente do Hudi.
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.
iceberg-defaults	Alterar os valores no arquivo iceberg-defaults.conf do Iceberg.
delta-defaults	Alterar os valores no arquivo delta-defaults.conf do Delta.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j2	Alterar as configurações de log4j2.properties do Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.

Classificações	Descrição
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.

Classificações	Descrição
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-lakeformation	Alterar os valores no arquivo lakeformation.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.

Classificações	Descrição
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
trino-log	Alterar os valores no arquivo log.properties do Trino.
trino-config	Alterar os valores no arquivo config.properties do Trino.
trino-password-authenticator	Alterar os valores no arquivo password-authenticator.properties do Trino.
trino-env	Alterar os valores no arquivo trino-env.sh do Trino.
trino-node	Alterar os valores no arquivo node.properties do Trino.
trino-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Trino.
trino-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Trino.
trino-connector-delta	Alterar os valores no arquivo delta.properties do Trino.
trino-connector-hive	Alterar os valores no arquivo hive.properties do Trino.

Classificações	Descrição
trino-exchange-manager	Alterar os valores no arquivo exchange-manager.properties do Trino.
trino-connector-iceberg	Alterar os valores no arquivo iceberg.properties do Trino.
trino-connector-hudi	Alterar os valores no arquivo hudi.properties do Trino.
trino-connector-jmx	Alterar os valores no arquivo jmx.properties do Trino.
trino-connector-kafka	Alterar os valores no arquivo kafka.properties do Trino.
trino-connector-localfile	Alterar os valores no arquivo localfile.properties do Trino.
trino-connector-memory	Alterar os valores no arquivo memory.properties do Trino.
trino-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Trino.
trino-connector-mysql	Alterar os valores no arquivo mysql.properties do Trino.
trino-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Trino.
trino-connector-raptor	Alterar os valores no arquivo raptor.properties do Trino.
trino-connector-redis	Alterar os valores no arquivo redis.properties do Trino.
trino-connector-redshift	Alterar os valores no arquivo redshift.properties do Trino.

Classificações	Descrição
trino-connector-tpch	Alterar os valores no arquivo tpch.properties do Trino.
trino-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Trino.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.
ranger-kms-logback	Alterar os valores no arquivo kms-logback.xml do Ranger KMS.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j2	Alterar os valores no arquivo log4j2.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.

Classificações	Descrição
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-logback	Altere os valores no ZooKeeper arquivo logback.xml do.
emr-metrics	Altere as configurações de métricas do EMR desse nó.

Log de alterações 7.1.0

Log de alterações da versão 7.1.0 e das notas da versão

Data	Event	Descrição
2024-05-10	Publicação de documentos	Notas da versão 7.1.0 do Amazon EMR publicadas pela primeira vez

Data	Event	Descrição
2024-05-08	Implantação concluída	Amazon EMR 7.1.0 totalmente implantado em todas as regiões compatíveis
2024-04-23	Lançamento inicial	A versão 7.1.0 do Amazon EMR foi implantada pela primeira vez em regiões comerciais iniciais

Versão 7.0.0 do Amazon EMR

Versões 7.0.0 da aplicação

Essa versão inclui os seguintes aplicativos: AmazonCloudWatchAgent, Delta, Flink, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, SqoopTensorFlow, Tez, TrinoZeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
AWS SDK for Java	2.25.70, 1.12.747	2.23.18, 1.12.705	2.23.18, 1.12.656	2.20.160-amzn-0, 1.12.569

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
Python	3,9, 3,11	3,9, 3,11	3,9, 3,11	3.9
Scala	2.12.18	2.12.17	2.12.17	2.12.17
AmazonCloudWatchAgent	1.300032.2	1.300032.2	1.300032.2	1.300031.1
Delta	3.2.0-amzn-0	3.1.0	3.0.0	3.0.0
Flink	1.18.1	1.18.1	1.18.1	1.18.0
Ganglia	-	-	-	-
HBase	2.4.17	2.4.17	2.4.17	2.4.17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.6	3.3.6	3.3.6	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-amzn-0	0.14.1-amzn-1	0.14.1-amzn-0	0.14.0-amzn-1
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.5.2-amzn-0	1.5.0-amzn-0	1.4.3-amzn-0	q1.4.2-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.7.1
MXNet	-	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
Phoenix	5.1.3	5.1.3	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,285	0,285	0,284	0,283
Spark	3.5.1	3.5.1	3.5.0	3.5.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.16.1	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	442	436	435	426
Zeppelin	0.11.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.9.1	3.9.1	3.9.1	3.5.10

Notas da versão 7.0.0

As notas da versão a seguir incluem informações sobre a versão 7.0.0 do Amazon EMR. As alterações são referentes à versão 6.15.0.

Novos recursos

- Atualizações da aplicação: Amazon EMR 7.0.0 application upgrades include Python 3.9, Spark 3.5, Flink 1.18, and Delta 3.0. Essa versão também adiciona suporte à aplicação do CloudWatch Agente da Amazon e remove o suporte ao Ganglia.
- Amazon Corretto 17: as versões 7.0 e superiores do Amazon EMR são fornecidas com o Amazon Corretto 17 (criado no OpenJDK) por padrão para aplicações compatíveis com o Corretto 17 (JDK 17), com exceção do Apache Livy. Para obter mais informações sobre as versões compatíveis do JDK para aplicações nessa versão, consulte Versões padrão do Java 7.0.0.
- Amazon Linux 2023 — Com a versão 7.0, os clusters do Amazon EMR agora são executados AL2 em 0.23 por padrão. Para obter informações sobre como isso afeta a versão padrão da AMI,

consulte Considerações sobre a atualização de software no Guia de gerenciamento do Amazon EMR. Observe também que AL2 023 removeu o Python 2.7, portanto, qualquer componente que exija Python agora deve ser escrito com o Python 3.

- S3 no Outposts com **`s3a`**: o Amazon EMR agora oferece suporte aos buckets do Amazon S3 no Outposts com o sistema de arquivos `s3a`. Para obter mais informações sobre o S3 no Outposts, consulte O que é o S3 no Outposts? no Guia do usuário do Amazon Simple Storage Service.

Problemas conhecidos

- Observe que você não pode executar mais de uma linha por vez se estiver usando PySpark as versões 3.10 ou posteriores do Python. Você deve executar uma linha de cada vez.

Alterações, melhorias e problemas resolvidos

- Para manter o estado de todas as instâncias em um cluster, o Amazon EMR se integra ao Apache YARN, Apache HDFS e Kerberos. Com a versão 7.0, aprimoramos essas integrações para remover de forma confiável o estado das instâncias que são encerradas devido ao ajuste de escala ou a outras operações. Isso é particularmente importante para clusters de longa duração que têm o ajuste de escala gerenciado habilitado, pois eles escalam com frequência e acumulam milhares de instâncias encerradas ao longo da vida útil.
- Essa versão melhora a configuração do Kerberos para incluir apenas o suporte a cifras baseadas em AES. O KDC do Kerberos com cifras que não baseadas em AES não é mais compatível com clusters do EMR executados nas versões 7.0.0 e superiores do Amazon EMR. As cifras baseadas em AES oferecem a maior segurança para os clusters.
- Como parte da migração do AWS SDK 2.x, o Amazon EMR 7.0 inclui uma atualização do conector Spark Kinesis para fins de compatibilidade. Essa atualização não está disponível na versão comunitária do Apache Spark. Se você usa o conector do Spark Kinesis de uma versão inferior à 7.0 do Amazon EMR, é necessário migrar os códigos da sua aplicação para execução no SDK 2.x antes de poder migrar suas workloads para o Amazon EMR 7.0. Para obter mais informações, consulte Migração do conector do Spark Kinesis para o SDK 2.x do Amazon EMR 7.0.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250527.1	6.1.134-152.225.amzn2023	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein), China (Pequim), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Leste dos EUA (Norte da Virgínia), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Malásia), Europa (Londres), Ásia-Pacífico (Melbourne), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (Oregon) AWS GovCloud , Oeste dos EUA (N. AWS GovCloud Califórnia), Ásia-Pacífico (Cingapura), Ásia-Pacífico

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Sydney), Israel (Tel Aviv), Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250512.(	6.1.134-152.225.am zn2023	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Hyderabad), Ásia-Pacífico (Jacarta)), Ásia-Pacífico (Malásia), Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Oriente Médio (Oriente Médio) (Oriente Médio (Oriente Médio)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Emirados Árabes Unidos), Canadá (Central) , Oeste do Canadá (Calgary), (Leste dos EUA), AWS GovCloud (AWS GovCloud Oeste dos EUA), China (Pequim), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250428.1	6.1.134-150.224.amzn2023	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Jacarta)), Ásia-Pacífico (Malásia), Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Oriente Médio (Oriente Médio) (Oriente Médio (Oriente Médio)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Emirados Árabes Unidos), Canadá (Central) , Oeste do Canadá (Calgary), (Leste dos EUA), AWS GovCloud (AWS GovCloud Oeste dos EUA), China (Pequim), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250414.(	6.1.132-147.221.am zn2023	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pacífico (Malásia),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250331.(	6.1.131-143.221.am zn2023	18 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pacífico (Malásia), Ásia-Pacífico (Tóquio)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250303.(	6.1.129-138.220.am zn2023	27 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pacífico (Malásia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250218.1	6.1.128-136.201.amzn2023	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderabad), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pacífico (Malásia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250211.(	6.1.127-135.201.am zn2023	26 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pacífico (Malásia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250123.4	6.1.124-134.200.amzn2023	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderabad), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250115.(	6.1.119-129.201.am zn2023	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 241121.(	6.1.115-126.197.am zn2023	12 de dezembro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 241031.(	6.1.112-124.190.am zn2023	15 de novembro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.5.2 241001.1	6.1.109-118.189.am zn2023	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.5.2 240819.(	6.1.102-111.182.am zn2023	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.5.2 240730.(	6.1.97-104.177.amzn2023	2 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.5.2 240722.(	6.1.97-104.177.amzn2023	24 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.5.2 240708.(	6.1.96-102.177.amzn2023	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), Oeste

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			do Canadá (Calgary) , (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China AWS GovCloud (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.3.2 240304.(	6.1.79-99.164.amzn2023	12 de março de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), Oeste

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			do Canadá (Calgary) , (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China AWS GovCloud (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.3.2 240219.(	6.1.77-99.164.amzn2023	1.º de março de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), Oeste

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			do Canadá (Calgary) , (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China AWS GovCloud (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.3.2 240205.(	6.1.75-99.163.amzn2023	19 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), Oeste

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			do Canadá (Calgary) , (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China AWS GovCloud (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.3.2 240122.(	6.1.72-96.166.amzn2023	5 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), Oeste

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			do Canadá (Calgary) , (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China AWS GovCloud (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.3.2 240108.(	6.1.72-96.166.amzn2023	24 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), Oeste

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			do Canadá (Calgary) , (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China AWS GovCloud (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.3.2 231211.4	6.1.66-91.160.amzn2023	19 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad) Abad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

Versões padrão do Java 7.0.0

O Amazon EMR 7.0 e versões superiores vêm com o Amazon Corretto 17 (criado no OpenJDK) por padrão para aplicações compatíveis com o Corretto 17 (JDK 17), com exceção do Apache Livy.

A tabela a seguir mostra as versões padrão do Java para aplicações no Amazon EMR 7.0.0. Se você quiser alterar a JVM padrão no cluster, siga as instruções em Configurar aplicações para usar uma Máquina Virtual Java específica para cada aplicação executada no cluster. Você só pode usar uma versão de runtime do Java para um cluster. O Amazon EMR não é compatível com a execução de diferentes nós ou aplicações em diferentes versões de runtime no mesmo cluster.

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
Delta	17, 11, 8
Flink	11, 8
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez	17, 11, 8
Trino	17
Zeppelin	8
Pig	8
Zookeeper	8

Versões de componente 7.0.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`*

começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
adot-java-agent	1.31.0	Um Java Agent que coleta métricas de daemons de aplicações.
delta	3.0.0	O Delta Lake é um formato de tabela aberto para grandes conjuntos de dados analíticos
delta-standalone-connectors	0.6.0	Os conectores do Delta oferecem runtimes diferente s para integrar o Delta Lake a mecanismos, como Flink, Hive e Presto.
emr-amazon-cloudwatch-agent	1.300031.1-amzn-0	Um aplicativo que coleta métricas internas em nível de sistema e métricas de aplicativ os personalizadas de instância s da Amazon. EC2
emr-ddb	5.2.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.9.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.13.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.7.0	Ambiente Conda para bloco de anotações do emr, que

Componente	Versão	Descrição
		inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.30.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	2.9.0	EMR S3Select Connector
emr-wal-cli	1.2.1	Cli usada para lista/exclusão de emrwal.
emrfs	2.61.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.18.0-amzn-0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.18.0-amzn-0	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
hadoop-client	3.3.6-amzn-2	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	3.3.6-amzn-2	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.3.6-amzn-2	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.3.6-amzn-2	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.

Componente	Versão	Descrição
hadoop-hdfs-journalnode	3.3.6-amzn-2	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.3.6-amzn-2	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.3.6-amzn-2	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	3.3.6-amzn-2	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.3.6-amzn-2	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	3.3.6-amzn-2	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	3.3.6-amzn-2	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	2.4.17-amzn-4	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.

Componente	Versão	Descrição
hbase-region-server	2.4.17-amzn-4	Serviço para atender a uma ou mais HBase regiões.
hbase-client	2.4.17-amzn-4	HBase cliente de linha de comando.
hbase-rest-server	2.4.17-amzn-4	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.4.17-amzn-4	Serviço que fornece um endpoint Thrift para. HBase
hbase-operator-tools	2.4.17-amzn-4	Ferramenta de reparo para HBase clusters Apache.
hcatalog-client	3.1.3-amzn-9	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.3-amzn-9	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	3.1.3-amzn-9	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.3-amzn-9	O cliente da linha de comando do Hive.
hive-hbase	3.1.3-amzn-9	Cliente do Hive-hbase.

Componente	Versão	Descrição
hive-metastore-server	3.1.3-amzn-9	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.3-amzn-9	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.14.0-amzn-1	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.14.0-amzn-1	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-trino	0.14.0-amzn-1	Biblioteca de pacotes para execução do Presto com o Hudi.
hudi-spark	0.14.0-amzn-1	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.11.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
iceberg	q1.4.2-amzn-0	Apache Iceberg é um formato de tabela aberta para conjuntos de dados analíticos imensos
jupyterhub	1.5.0	Servidor de vários usuários para blocos de anotações Jupyter

Componente	Versão	Descrição
livy-server	0.7.1-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mxnet	1.9.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.
nvidia-cuda	11.8.0	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.7.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	5.1.3	As bibliotecas do phoenix para servidor e cliente
phoenix-connectors	5.1.3	Conectores do Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica

Componente	Versão	Descrição
presto-coordinator	0.283-amzn-1	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.283-amzn-1	O serviço que executa partes de uma consulta.
presto-client	0.283-amzn-1	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
trino-coordinator	426-amzn-1	Serviço que aceita consultas e gerencia a execução de consultas entre trino-workers.
trino-worker	426-amzn-1	O serviço que executa partes de uma consulta.
trino-client	426-amzn-1	Cliente de linha de comando do Trino instalado em mestres de espera de um cluster de alta disponibilidade (HA) em que o servidor do Trino não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.3.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger

Componente	Versão	Descrição
spark-client	3.5.0-amzn-0	Os clientes da linha de comando do Spark.
spark-history-server	3.5.0-amzn-0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.5.0-amzn-0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.5.0-amzn-0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
spark-rapids	23.10.0-amzn-0	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.11.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.10.2-amzn-7	O aplicativo e as bibliotecas do YARN para Tez.
tez-on-worker	0.10.2-amzn-7	A aplicação e as bibliotecas do YARN nós de processam ento.
webserver	2.4.58	O servidor Apache HTTP.

Componente	Versão	Descrição
zeppelin-server	0.10.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.5.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.5.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 7.0.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-7.0.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.

Classificações	Descrição
docker-conf	Alterar as configurações relacionadas ao docker.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-session	Altere as configurações do Flink log4j-ses sion.properties para a sessão. Kubernetes/Yarn
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.

Classificações	Descrição
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-env	Alterar os valores no ambiente do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.

Classificações	Descrição
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.proper ties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.x ml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-java-home	Alterar a página inicial do KMS do Hadoop
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properti es do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.

Classificações	Descrição
hudi-env	Altere os valores no ambiente do Hudi.
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.
iceberg-defaults	Alterar os valores no arquivo iceberg-defaults.conf do Iceberg.
delta-defaults	Alterar os valores no arquivo delta-defaults.conf do Delta.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j2	Alterar as configurações de log4j2.properties do Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.

Classificações	Descrição
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.

Classificações	Descrição
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-lakeformation	Alterar os valores no arquivo lakeformation.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.

Classificações	Descrição
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
trino-log	Alterar os valores no arquivo log.properties do Trino.
trino-config	Alterar os valores no arquivo config.properties do Trino.
trino-password-authenticator	Alterar os valores no arquivo password-authenticator.properties do Trino.
trino-env	Alterar os valores no arquivo trino-env.sh do Trino.
trino-node	Alterar os valores no arquivo node.properties do Trino.
trino-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Trino.
trino-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Trino.
trino-connector-delta	Alterar os valores no arquivo delta.properties do Trino.
trino-connector-hive	Alterar os valores no arquivo hive.properties do Trino.

Classificações	Descrição
trino-exchange-manager	Alterar os valores no arquivo exchange-manager.properties do Trino.
trino-connector-iceberg	Alterar os valores no arquivo iceberg.properties do Trino.
trino-connector-hudi	Alterar os valores no arquivo hudi.properties do Trino.
trino-connector-jmx	Alterar os valores no arquivo jmx.properties do Trino.
trino-connector-kafka	Alterar os valores no arquivo kafka.properties do Trino.
trino-connector-localfile	Alterar os valores no arquivo localfile.properties do Trino.
trino-connector-memory	Alterar os valores no arquivo memory.properties do Trino.
trino-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Trino.
trino-connector-mysql	Alterar os valores no arquivo mysql.properties do Trino.
trino-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Trino.
trino-connector-raptor	Alterar os valores no arquivo raptor.properties do Trino.
trino-connector-redis	Alterar os valores no arquivo redis.properties do Trino.
trino-connector-redshift	Alterar os valores no arquivo redshift.properties do Trino.

Classificações	Descrição
trino-connector-tpch	Alterar os valores no arquivo tpch.properties do Trino.
trino-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Trino.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.
ranger-kms-logback	Alterar os valores no arquivo kms-logback.xml do Ranger KMS.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j2	Alterar os valores no arquivo log4j2.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.

Classificações	Descrição
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Log de alterações 7.0.0

Log de alterações da versão 7.0.0 e das notas da versão

Data	Event	Descrição
2023-12-29	Implantação concluída	Amazon EMR 7.0.0 totalmente implantado em todas as regiões compatíveis

Data	Event	Descrição
2023-12-28	Publicação de documentos	Notas da versão 7.0.0 do Amazon EMR publicadas pela primeira vez
2023-12-19	Lançamento inicial	A versão 7.0.0 do Amazon EMR foi implantada pela primeira vez em regiões comerciais iniciais

Versões de lançamento 6.x do Amazon EMR

Esta seção contém versões de aplicações, notas da versão, versões de componentes e classificações de configuração disponíveis em cada versão de lançamento 6.x do Amazon EMR.

Ao executar um cluster, você pode escolher entre várias versões do Amazon EMR. Isso permite que você teste e use versões de aplicações que atendem aos seus requisitos de compatibilidade. Você especifica o número da versão com o rótulo da versão. Os rótulos de versão estão no formato `emr-x.x.x`. Por exemplo, `.emr-7.9.0`

Novas versões do Amazon EMR são disponibilizadas em diferentes regiões durante um período de vários anos, começando com a primeira região na data da versão inicial. A versão mais recente pode não estar disponível em sua região durante esse período.

Para obter uma tabela abrangente das versões de aplicações em cada versão 6.x do Amazon EMR, consulte Versões de aplicações nas versões 6.x do Amazon EMR.

Tópicos

- Versões de aplicações nas versões 6.x do Amazon EMR
- Versão 6.15.0 do Amazon EMR
- Versão 6.14.0 do Amazon EMR
- Versão 6.13.0 do Amazon EMR
- Versão 6.12.0 do Amazon EMR
- Versão 6.11.1 do Amazon EMR
- Versão 6.11.0 do Amazon EMR

- Versão 6.10.1 do Amazon EMR
- Versão 6.10.0 do Amazon EMR
- Versão 6.9.1 do Amazon EMR
- Versão 6.9.0 do Amazon EMR
- Versão 6.8.1 do Amazon EMR
- Versão 6.8.0 do Amazon EMR
- Versão 6.7.0 do Amazon EMR
- Versão 6.6.0 do Amazon EMR
- Versão 6.5.0 do Amazon EMR
- Versão 6.4.0 do Amazon EMR
- Versão 6.3.1 do Amazon EMR
- Versão 6.3.0 do Amazon EMR
- Versão 6.2.1 do Amazon EMR
- Versão 6.2.0 do Amazon EMR
- Versão 6.1.1 do Amazon EMR
- Versão 6.1.0 do Amazon EMR
- Versão 6.0.1 do Amazon EMR
- Versão 6.0.0 do Amazon EMR

Versões de aplicações nas versões 6.x do Amazon EMR

Para obter uma tabela abrangente que liste as versões das aplicações disponíveis em cada versão 6.x do Amazon EMR, abra as versões de aplicações nas versões 6.x do Amazon EMR no seu navegador.

Versão 6.15.0 do Amazon EMR

Versões 6.15.0 da aplicação

Essa versão inclui os seguintes aplicativos: Delta, Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, SqoopTensorFlow, Tez, TrinoZeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-6.15.0	emr-6.14.0	emr-6.13.0	emr-6.12.0
AWS SDK for Java	2.20.160-amzn-0, 1.12.569	1.12.543	1.12.513	1.12.490
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.17	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.4.0	2.4.0	2.4.0	2.4.0
Flink	1.17.1	1.17.1	1.17.0	1.17.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.17	2.4.17	2.4.17	2.4.17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.6	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3

	emr-6.15.0	emr-6.14.0	emr-6.13.0	emr-6.12.0
Hudi	0.14.0-amzn-0	0.13.1-amzn-2	0.13.1-amzn-1	0.13.1-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.4.0-amzn-0	1.3.1-amzn-0	1.3.0-amzn-1	1.3.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,283	0,281	0,281	0,281
Spark	3.4.1	3.4.1	3.4.1	3.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	426	422	414	414
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

Notas da versão 6.15.0

As notas da versão a seguir incluem informações sobre a versão 6.15.0 do Amazon EMR. As alterações são referentes à versão 6.14.0. Para obter informações sobre o cronograma da versão, consulte o Log de alterações 6.15.0.

Novos recursos

- Atualizações da aplicação: Amazon EMR 6.15.0 application upgrades include Apache Hadoop 3.3.6, Apache Hudi 0.14.0-amzn-0, Iceberg 1.4.0-amzn-0, and Trino 426.
- Lançamentos mais rápidos para clusters do EMR que são executados em EC2 — Agora é até 35% mais rápido lançar um Amazon EMR em um cluster. EC2 Com essa melhoria, a maioria dos clientes pode lançar seus clusters em cinco minutos ou menos.
- CodeWhisperer para o EMR Studio — Agora você pode usar a Amazon com o CodeWhisperer Amazon EMR Studio para obter recomendações em tempo real à medida que você escreve código. JupyterLab CodeWhisperer pode concluir seus comentários, concluir linhas únicas de código, fazer line-by-line recomendações e gerar funções totalmente formadas.
- Tempos de reinicialização de trabalhos mais rápidos com o Flink: com o Amazon EMR 6.15.0 e superior, vários novos mecanismos estão disponíveis para o Apache Flink a fim de melhorar o tempo de reinicialização de trabalhos durante as operações de recuperação ou escalabilidade de tarefas. Isso otimiza a velocidade de recuperação e reinicialização dos gráficos de execução para melhorar a estabilidade do trabalho.
- Controle de acesso detalhado e em nível de tabela para formatos de tabela aberta — Com o Amazon EMR 6.15.0 e superior, quando você executa trabalhos do Spark no Amazon EMR em EC2 clusters que acessam dados no AWS Glue Data Catalog, você pode AWS Lake Formation usar para aplicar permissões em nível de tabela, linha, coluna e célula em tabelas baseadas em Hudi, Iceberg ou Delta Lake.
- Atualização do Hadoop: o Amazon EMR 6.15.0 inclui uma atualização do Apache Hadoop para a versão 3.3.6. O Hadoop 3.3.6 era a versão mais recente na época da implantação do Amazon EMR 6.15, lançada pela Apache em junho de 2023. Versões anteriores do Amazon EMR (6.9.0 até 6.14.x) usavam o Hadoop 3.3.3.

  A atualização inclui centenas de melhorias e correções, além de recursos com parâmetros de nós de dados reconfiguráveis, a opção `DFSAdmin` para iniciar operações de reconfiguração em massa em todos os nós de dados ativos e uma API vetorizada que permite aos leitores exigentes especificar vários intervalos de leitura. O Hadoop 3.3.6 também adiciona suporte ao HDFS APIs e semântica ao seu registro de gravação antecipada (WAL), para que ele possa ser executado

em outras implementações de sistemas de armazenamento. HBase Para obter mais informações, consulte os logs de alterações das versões 3.3.4, 3.3.5 e 3.3.6 na documentação do Apache Hadoop.

- Suporte ao AWS SDK for Java, versão 2 - Os aplicativos Amazon EMR 6.15.0 podem usar o SDK for AWS Java nas versões 1.12.569 ou 2.20.160 se o aplicativo suportar a versão 2. O AWS SDK for Java 2.x é uma grande reescrita da base de código da versão 1.x. Ele foi criado com base no Java 8+ e adiciona vários recursos frequentemente solicitados. Entre eles, suporte para E/S sem bloqueio e capacidade de conectar uma implementação HTTP diferente no runtime. Para obter mais informações, incluindo um Guia de migração do SDK para Java da v1 à v2, consulte o guia AWS SDK para Java, versão 2.

Alterações, melhorias e problemas resolvidos

- Para melhorar seus clusters de alta disponibilidade do EMR, essa versão permite a conectividade com daemons do Amazon EMR em um host local que usa endpoints IPv6.
- Esta versão habilita o TLS 1.2 para comunicação com ZooKeeper provisionados em todos os nós primários do seu cluster de alta disponibilidade.
- Esta versão melhora o gerenciamento dos arquivos de log de ZooKeeper transações que são mantidos nos nós primários para minimizar os cenários em que os arquivos de log ultrapassam os limites e interrompem as operações do cluster.
- Essa versão torna a comunicação entre nós mais resiliente para clusters de alta disponibilidade do EMR. Essa melhoria reduz a chance de falhas na ação de bootstrap ou na inicialização do cluster.
- O Tez no Amazon EMR 6.15.0 introduz configurações que você pode especificar para abrir de forma assíncrona as divisões de entrada em uma divisão agrupada do Tez. Isso resulta em uma performance mais rápida das consultas de leitura quando há um grande número de divisões de entrada em uma única divisão agrupada do Tez. Para obter mais informações, consulte a Abertura assíncrona da divisão do Tez.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Malásia), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Malásia), Europa (Londres), Ásia-Pacífico (Melbourne), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (AWS GovCloud AWS GovCloud Oregon), México (Central) , Oeste dos EUA (Norte da

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Califórnia), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv), Ásia-Pacífico (Taipei), Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Hyderabad), Ásia-Pacífico (Jacarta) Arta), Ásia-Pacífico (Malásia), Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Malásia) , Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (EAU), Canadá (Central) , Oeste do Canadá

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Calgary), (Leste dos EUA), AWS GovCloud (Oeste dos EUA), China (AWS GovCloud Pequim), China (Ningxia), Israel (Tel Aviv), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Centro Leste (Emirados Árabes Unidos), Ásia-Pacífico (Melbourne),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Israel (Tel Aviv), Oeste do Canadá (Calgary), Europa (Zurique), Ásia-Pacífico (Malásia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 001.0	4.14.352	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 816.0	4.14.350	21 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 809.0	4.14.349	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 719.0	4.14.348	25 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 223.0	4.14.336	8 de março de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 131.0	4.14.336	14 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 124.0	4.14.336	7 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 109.0	4.14.334	24 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 218.0	4.14.330	2 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 206.0	4.14.330	22 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 116.0	4.14.328	11 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 101.0	4.14.327	13 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

Versões padrão do Java 6.15.0

As versões 6.12.0 e superiores do Amazon EMR são compatíveis com todas as aplicações que têm o Amazon Corretto 8 por padrão, exceto o Trino. Para Trino, o Amazon EMR é compatível com o Amazon Corretto 17 por padrão, desde a versão 6.9.0 do Amazon EMR. O Amazon EMR também é compatível com algumas aplicações com o Amazon Corretto 11 e 17. Essas aplicações estão listadas na tabela a seguir. Se você quiser alterar a JVM padrão no cluster, siga as instruções em Configurar aplicações para usar uma Máquina Virtual Java específica para cada aplicação executada no cluster. Você só pode usar uma versão de runtime do Java para um cluster. O Amazon EMR não é compatível com a execução de diferentes nós ou aplicações em diferentes versões de runtime no mesmo cluster.

Embora o Amazon EMR seja compatível com o Amazon Corretto 11 e 17 no Apache Spark, no Apache Hadoop e no Apache Hive, a performance de algumas workloads pode diminuir quando você usa estas versões do Corretto. Recomendamos que você teste suas workloads antes de alterar os padrões.

A seguinte tabela mostra as versões padrão do Java para aplicações no Amazon EMR 6.15.0:

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
Delta	17, 11, 8
Flink	11, 8
Ganglia	8

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez	17, 11, 8
Trino	17
Zeppelin	8
Pig	8
Zookeeper	8

Versões do componente 6.15.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.2	SDK do Amazon SageMaker Spark
delta	2.4.0	O Delta Lake é um formato de tabela aberto para grandes conjuntos de dados analíticos
delta-standalone-connectors	0.6.0	Os conectores do Delta oferecem runtimes diferente s para integrar o Delta Lake a mecanismos, como Flink, Hive e Presto.
emr-ddb	5.2.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.8.0	Bibliotecas convenientes para o ecossistema do Hadoop.

Componente	Versão	Descrição
emr-kinesis	3.12.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.7.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.29.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	2.8.0	EMR S3Select Connector
emr-wal-cli	1.2.0	Cli usada para lista/exclusão de emrwal.
emrfs	2.60,0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.17.1-amzn-1	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.17.1-amzn-1	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	3.3.6-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	3.3.6-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.3.6-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.3.6-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	3.3.6-amzn-1	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.3.6-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.3.6-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider

Componente	Versão	Descrição
hadoop-mapred	3.3.6-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.3.6-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	3.3.6-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hadoop-yarn-timeline-server	3.3.6-amzn-1	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	2.4.17-amzn-3	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	2.4.17-amzn-3	Serviço para atender a uma ou mais HBase regiões.
hbase-client	2.4.17-amzn-3	HBase cliente de linha de comando.
hbase-rest-server	2.4.17-amzn-3	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.4.17-amzn-3	Serviço que fornece um endpoint Thrift para. HBase

Componente	Versão	Descrição
hbase-operator-tools	2.4.17-amzn-3	Ferramenta de reparo para HBase clusters Apache.
hcatalog-client	3.1.3-amzn-8	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.3-amzn-8	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	3.1.3-amzn-8	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.3-amzn-8	O cliente da linha de comando do Hive.
hive-hbase	3.1.3-amzn-8	Cliente do Hive-hbase.
hive-metastore-server	3.1.3-amzn-8	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.3-amzn-8	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.14.0-amzn-0	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.

Componente	Versão	Descrição
hudi-presto	0.14.0-amzn-0	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-trino	0.14.0-amzn-0	Biblioteca de pacotes para execução do Presto com o Hudi.
hudi-spark	0.14.0-amzn-0	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.11.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
iceberg	1.4.0-amzn-0	Apache Iceberg é um formato de tabela aberta para conjuntos de dados analíticos imensos
jupyterhub	1.5.0	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.1-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mxnet	1.9.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.

Componente	Versão	Descrição
nvidia-cuda	11.8.0	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.7.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	5.1.3	As bibliotecas do phoenix para servidor e cliente
phoenix-connectors	5.1.3	Conectores do Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.283-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.283-amzn-0	O serviço que executa partes de uma consulta.

Componente	Versão	Descrição
presto-client	0.283-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
trino-coordinator	426-amzn-0	Serviço que aceita consultas e gerencia a execução de consultas entre trino-workers.
trino-worker	426-amzn-0	O serviço que executa partes de uma consulta.
trino-client	426-amzn-0	Cliente de linha de comando do Trino instalado em mestres de espera de um cluster de alta disponibilidade (HA) em que o servidor do Trino não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.0.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	3.4.1-amzn-2	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	3.4.1-amzn-2	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.4.1-amzn-2	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.4.1-amzn-2	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
spark-rapids	23.08.1-amzn-0	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.11.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.10.2-amzn-6	O aplicativo e as bibliotecas do YARN para Tez.
tez-on-worker	0.10.2-amzn-6	A aplicação e as bibliotecas do YARN nós de processam ento.
webserver	2.4.41+	O servidor Apache HTTP.
zeppelin-server	0.10.1	O notebook baseado na web que permite um data analytics interativo.

Componente	Versão	Descrição
zookeeper-server	3.5.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.5.10	ZooKeeper cliente de linha de comando.

Classificações da configuração 6.15.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-6.15.0

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Classificações	Descrição	Ações de reconfiguração
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.

Classificações	Descrição	Ações de reconfiguração
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Alterar a página inicial do KMS do Hadoop	Not available.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.

Classificações	Descrição	Ações de reconfiguração
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.	Not available.
iceberg-defaults	Alterar os valores no arquivo iceberg-defaults.conf do Iceberg.	Not available.
delta-defaults	Alterar os valores no arquivo delta-defaults.conf do Delta.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j2	Alterar as configurações de log4j2.properties do Livy.	Restarts Livy Server.

Classificações	Descrição	Ações de reconfiguração
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.

Classificações	Descrição	Ações de reconfiguração
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-lakeformation	Alterar os valores no arquivo lakeformation.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
trino-log	Alterar os valores no arquivo log.properties do Trino.	Restarts Trino-Server (for Trino)
trino-config	Alterar os valores no arquivo config.properties do Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Alterar os valores no arquivo password-authenticator.properties do Trino.	Restarts Trino-Server (for Trino)
trino-env	Alterar os valores no arquivo trino-env.sh do Trino.	Restarts Trino-Server (for Trino)
trino-node	Alterar os valores no arquivo node.properties do Trino.	Not available.
trino-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Trino.	Not available.
trino-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Trino.	Not available.
trino-connector-delta	Alterar os valores no arquivo delta.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Alterar os valores no arquivo hive.properties do Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Alterar os valores no arquivo exchange-manager.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Alterar os valores no arquivo iceberg.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Alterar os valores no arquivo hudi.properties do Trino.	Restarts Trino-Server (for Trino)

Classificações	Descrição	Ações de reconfiguração
trino-connector-jmx	Alterar os valores no arquivo jmx.properties do Trino.	Not available.
trino-connector-kafka	Alterar os valores no arquivo kafka.properties do Trino.	Not available.
trino-connector-localfile	Alterar os valores no arquivo localfile.properties do Trino.	Not available.
trino-connector-memory	Alterar os valores no arquivo memory.properties do Trino.	Not available.
trino-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Trino.	Not available.
trino-connector-mysql	Alterar os valores no arquivo mysql.properties do Trino.	Not available.
trino-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Trino.	Not available.
trino-connector-raptor	Alterar os valores no arquivo raptor.properties do Trino.	Not available.
trino-connector-redis	Alterar os valores no arquivo redis.properties do Trino.	Not available.
trino-connector-redshift	Alterar os valores no arquivo redshift.properties do Trino.	Not available.
trino-connector-tpch	Alterar os valores no arquivo tpch.properties do Trino.	Not available.
trino-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Trino.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Alterar os valores no arquivo kms-logback.xml do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j2	Alterar os valores no arquivo log4j2.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie and HiveServe r2.
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.

Classificações	Descrição	Ações de reconfiguração
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Log de alterações 6.15.0

Log de alterações da versão 6.15.0 e das notas da versão

Data	Event	Descrição
2023-11-17	Publicação de documentos	Notas da versão 6.15.0 do Amazon EMR publicadas pela primeira vez
2023-11-17	Implantação concluída	Amazon EMR 6.15.0 totalmente implantado em todas as regiões compatíveis
2023-11-13	Lançamento inicial	A versão 6.15.0 do Amazon EMR foi implantada pela primeira vez em regiões comerciais iniciais

Versão 6.14.0 do Amazon EMR

Versões da aplicação 6.14.0

Essa versão inclui os seguintes aplicativos: Delta, Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, SqoopTensorFlow, Tez, TrinoZeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-6.14.0	emr-6.13.0	emr-6.12.0	emr-6.11.1
AWS SDK for Java	1.12.543	1.12.513	1.12.490	1.12.446
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.4.0	2.4.0	2.4.0	2.2.0
Flink	1.17.1	1.17.0	1.17.0	1.16.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.17	2.4.17	2.4.17	2.4.15
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.13.1-amzn-2	0.13.1-amzn-1	0.13.1-amzn-0	0.13.0-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.3.1-amzn-0	1.3.0-amzn-1	1.3.0-amzn-0	1.2.0-amzn-0

	emr-6.14.0	emr-6.13.0	emr-6.12.0	emr-6.11.1
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.3	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,281	0,281	0,281	0,279
Spark	3.4.1	3.4.1	3.4.0	3.3.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	422	414	414	410
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

Notas da versão 6.14.0

As notas da versão a seguir incluem informações para a versão 6.14.0 do Amazon EMR. As alterações são referentes à versão 6.13.0. Para obter informações sobre o cronograma da versão, consulte o Log de alterações 6.14.0.

Novos recursos

- Amazon EMR 6.14.0 supports Apache Spark 3.4.1, Apache Spark RAPIDS 23.06.0-amzn-2, Flink 1.17.1, Iceberg 1.3.1, and Trino 422.
- O Ajuste de Escala Gerenciado do Amazon EMR agora está disponível na região `ap-southeast-3` Ásia-Pacífico (Jacarta) para clusters criados com o Amazon EMR 6.14.0 e superior.

Alterações, melhorias e problemas resolvidos

- A versão 6.14.0 otimiza o gerenciamento de logs com o Amazon EMR em execução na Amazon. EC2 Como resultado, é possível ver uma pequena redução nos custos de armazenamento dos logs do cluster.
- A versão 6.14.0 melhora o fluxo de trabalho de ajuste de escala para considerar diferentes instâncias principais que têm variação substancial no tamanho dos volumes do Amazon EBS. Essa melhoria se aplica somente aos nós centrais; as operações de redução dos nós de tarefas não são afetadas.
- A versão 6.14.0 melhora a forma como o Amazon EMR interage com aplicações de código aberto, como. Apache Hadoop YARN ResourceManager and HDFS NameNode Essa melhoria reduz o risco de atrasos operacionais com o escalonamento do cluster e atenua as falhas de inicialização que ocorrem devido a problemas de conectividade com os aplicações de código aberto.
- A versão 6.14.0 otimiza a instalação da aplicação na inicialização do cluster. Isso melhora os tempos de inicialização do cluster para determinadas combinações de aplicações do Amazon EMR.
- A versão 6.14.0 corrige um problema em que as operações de redução de escala do cluster podem parar quando um cluster executado em uma VPC com um domínio personalizado é reiniciado no nó central ou da tarefa.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Malásia), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Malásia), Europa (Londres), Ásia-Pacífico (Melbourne), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (AWS GovCloud AWS GovCloud Oregon), México (Central) , Oeste dos EUA (Norte da

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Califórnia), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv), Ásia-Pacífico (Taipei), Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Hyderabad), Ásia-Pacífico (Jacarta) Arta), Ásia-Pacífico (Malásia), Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Malásia) , Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (EAU), Canadá (Central) , Oeste do Canadá

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Calgary), (Leste dos EUA), AWS GovCloud (Oeste dos EUA), China (AWS GovCloud Pequim), China (Ningxia), Israel (Tel Aviv), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Centro Leste (Emirados Árabes Unidos), Ásia-Pacífico (Melbourne),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Israel (Tel Aviv), Oeste do Canadá (Calgary), Europa (Zurique), Ásia-Pacífico (Malásia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 223.0	4.14.336	8 de março de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 131.0	4.14.336	14 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 124.0	4.14.336	7 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 109.0	4.14.334	24 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 218.0	4.14.330	2 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 206.0	4.14.330	22 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 116.0	4.14.328	11 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 101.0	4.14.327	17 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 906.0	4.14.322	11 de setembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

Versões padrão do Java 6.14.0

As versões 6.12.0 e superiores do Amazon EMR são compatíveis com todas as aplicações que têm o Amazon Corretto 8 por padrão, exceto o Trino. Para Trino, o Amazon EMR é compatível com o Amazon Corretto 17 por padrão, desde a versão 6.9.0 do Amazon EMR. O Amazon EMR também é compatível com algumas aplicações com o Amazon Corretto 11 e 17. Essas aplicações estão listadas na tabela a seguir. Se você quiser alterar a JVM padrão no cluster, siga as instruções em Configurar aplicações para usar uma Máquina Virtual Java específica para cada aplicação executada no cluster. Você só pode usar uma versão de runtime do Java para um cluster. O Amazon EMR não é compatível com a execução de diferentes nós ou aplicações em diferentes versões de runtime no mesmo cluster.

Embora o Amazon EMR seja compatível com o Amazon Corretto 11 e 17 no Apache Spark, no Apache Hadoop e no Apache Hive, a performance de algumas workloads pode diminuir quando você usa estas versões do Corretto. Recomendamos que você teste suas workloads antes de alterar os padrões.

A seguinte tabela mostra as versões padrão do Java para aplicações no Amazon EMR 6.14.0:

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
Delta	17, 11, 8
Flink	11, 8
Ganglia	8
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11, 8

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez	17, 11, 8
Trino	17
Zeppelin	8
Pig	8
Zookeeper	8

Versões do componente 6.14.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado

`myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.2	SDK do Amazon SageMaker Spark
delta	2.4.0	O Delta Lake é um formato de tabela aberto para grandes conjuntos de dados analíticos
delta-standalone-connectors	0.6.0	Os conectores do Delta oferecem runtimes diferente s para integrar o Delta Lake a mecanismos, como Flink, Hive e Presto.
emr-ddb	5.1.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.7.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.11.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.7.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.28.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	2.7.0	EMR S3Select Connector

Componente	Versão	Descrição
emr-wal-cli	1.1.0	Cli usada para lista/exclusão de emrwal.
emrfs	2.59.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.17.1-amzn-0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.17.1-amzn-0	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	3.3.3-amzn-6	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.

Componente	Versão	Descrição
hadoop-hdfs-datanode	3.3.3-amzn-6	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.3.3-amzn-6	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.3.3-amzn-6	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	3.3.3-amzn-6	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.3.3-amzn-6	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.3.3-amzn-6	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	3.3.3-amzn-6	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.3.3-amzn-6	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	3.3.3-amzn-6	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.

Componente	Versão	Descrição
hadoop-yarn-timeline-server	3.3.3-amzn-6	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	2.4.17-amzn-2	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	2.4.17-amzn-2	Serviço para atender a uma ou mais HBase regiões.
hbase-client	2.4.17-amzn-2	HBase cliente de linha de comando.
hbase-rest-server	2.4.17-amzn-2	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.4.17-amzn-2	Serviço que fornece um endpoint Thrift para. HBase
hbase-operator-tools	2.4.17-amzn-2	Ferramenta de reparo para HBase clusters Apache.
hcatalog-client	3.1.3-amzn-7	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.3-amzn-7	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	3.1.3-amzn-7	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.3-amzn-7	O cliente da linha de comando do Hive.
hive-hbase	3.1.3-amzn-7	Cliente do Hive-hbase.
hive-metastore-server	3.1.3-amzn-7	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.3-amzn-7	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.13.1-amzn-2	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.13.1-amzn-2	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-trino	0.13.1-amzn-2	Biblioteca de pacotes para execução do Presto com o Hudi.
hudi-spark	0.13.1-amzn-2	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.11.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop

Componente	Versão	Descrição
iceberg	1.3.1-amzn-0	Apache Iceberg é um formato de tabela aberta para conjuntos de dados analíticos imensos
jupyterhub	1.5.0	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.1-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mxnet	1.9.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.
nvidia-cuda	11.8.0	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.7.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	5.1.3	As bibliotecas do phoenix para servidor e cliente

Componente	Versão	Descrição
phoenix-connectors	5.1.3	Conectores do Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.281-amzn-2	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.281-amzn-2	O serviço que executa partes de uma consulta.
presto-client	0.281-amzn-2	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
trino-coordinator	422-amzn-0	Serviço que aceita consultas e gerencia a execução de consultas entre trino-workers.
trino-worker	422-amzn-0	O serviço que executa partes de uma consulta.

Componente	Versão	Descrição
trino-client	422-amzn-0	Cliente de linha de comando do Trino instalado em mestres de espera de um cluster de alta disponibilidade (HA) em que o servidor do Trino não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.0.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	3.4.1-amzn-1	Os clientes da linha de comando do Spark.
spark-history-server	3.4.1-amzn-1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.4.1-amzn-1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.4.1-amzn-1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
spark-rapids	23.06.0-amzn-2	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs

Componente	Versão	Descrição
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.11.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.10.2-amzn-5	O aplicativo e as bibliotecas do YARN para Tez.
tez-on-worker	0.10.2-amzn-5	A aplicação e as bibliotecas do YARN nós de processam ento.
webserver	2.4.41+	O servidor Apache HTTP.
zeppelin-server	0.10.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.5.10	O serviço centralizado de manutenção de informações de configuração, nomenclat ura, fornecimento de sincroniz ação distribuída, e fornecime nto de serviços de grupo.
zookeeper-client	3.5.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 6.14.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-6.14.0

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.

Classificações	Descrição	Ações de reconfiguração
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.

Classificações	Descrição	Ações de reconfiguração
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.

Classificações	Descrição	Ações de reconfiguração
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classificações	Descrição	Ações de reconfiguração
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.

Classificações	Descrição	Ações de reconfiguração
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Alterar a página inicial do KMS do Hadoop	Not available.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.	Not available.
iceberg-defaults	Alterar os valores no arquivo iceberg-defaults.conf do Iceberg.	Not available.
delta-defaults	Alterar os valores no arquivo delta-defaults.conf do Delta.	Not available.

Classificações	Descrição	Ações de reconfiguração
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j2	Alterar as configurações de log4j2.properties do Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.

Classificações	Descrição	Ações de reconfiguração
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-lakeformation	Alterar os valores no arquivo lakeformation.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
trino-log	Alterar os valores no arquivo log.properties do Trino.	Restarts Trino-Server (for Trino)
trino-config	Alterar os valores no arquivo config.properties do Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Alterar os valores no arquivo password-authenticator.properties do Trino.	Restarts Trino-Server (for Trino)
trino-env	Alterar os valores no arquivo trino-env.sh do Trino.	Restarts Trino-Server (for Trino)

Classificações	Descrição	Ações de reconfiguração
trino-node	Alterar os valores no arquivo node.properties do Trino.	Not available.
trino-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Trino.	Not available.
trino-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Trino.	Not available.
trino-connector-delta	Alterar os valores no arquivo delta.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Alterar os valores no arquivo hive.properties do Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Alterar os valores no arquivo exchange-manager.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Alterar os valores no arquivo iceberg.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Alterar os valores no arquivo hudi.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Alterar os valores no arquivo jmx.properties do Trino.	Not available.
trino-connector-kafka	Alterar os valores no arquivo kafka.properties do Trino.	Not available.
trino-connector-localfile	Alterar os valores no arquivo localfile.properties do Trino.	Not available.
trino-connector-memory	Alterar os valores no arquivo memory.properties do Trino.	Not available.

Classificações	Descrição	Ações de reconfiguração
trino-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Trino.	Not available.
trino-connector-mysql	Alterar os valores no arquivo mysql.properties do Trino.	Not available.
trino-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Trino.	Not available.
trino-connector-raptor	Alterar os valores no arquivo raptor.properties do Trino.	Not available.
trino-connector-redis	Alterar os valores no arquivo redis.properties do Trino.	Not available.
trino-connector-redshift	Alterar os valores no arquivo redshift.properties do Trino.	Not available.
trino-connector-tpch	Alterar os valores no arquivo tpch.properties do Trino.	Not available.
trino-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Trino.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Alterar os valores no arquivo kms-logback.xml do Ranger KMS.	Not available.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
spark	As configurações administr adas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j2	Alterar os valores no arquivo log4j2.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie and HiveServe r2.

Classificações	Descrição	Ações de reconfiguração
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Log de alterações 6.14.0

Log de alterações da versão 6.14.0 e das notas da versão

Data	Event	Descrição
2023-11-02	Implantação concluída	Amazon EMR 6.14.0 totalmente implantado em todas as regiões compatíveis
2023-10-10	Publicação de documentos	Notas da versão 6.14.0 do Amazon EMR publicadas pela primeira vez
2023-10-04	Lançamento inicial	A versão 6.14.0 do Amazon EMR foi implantada pela primeira vez em regiões comerciais iniciais

Versão 6.13.0 do Amazon EMR

Versões da aplicação 6.13.0

Essa versão inclui os seguintes aplicativos: Delta, Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, SqoopTensorFlow, Tez, TrinoZeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-6.13.0	emr-6.12.0	emr-6.11.1	emr-6.11.0
AWS SDK for Java	1.12.513	1.12.490	1.12.446	1.12.446
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.4.0	2.4.0	2.2.0	2.2.0
Flink	1.17.0	1.17.0	1.16.0	1.16.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.17	2.4.17	2.4.15	2.4.15
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.13.1-amzn-1	0.13.1-amzn-0	0.13.0-amzn-0	0.13.0-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.3.0-amzn-1	1.3.0-amzn-0	1.2.0-amzn-0	1.2.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1

	emr-6.13.0	emr-6.12.0	emr-6.11.1	emr-6.11.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,281	0,281	0,279	0,279
Spark	3.4.1	3.4.0	3.3.2	3.3.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	414	414	410	410
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

Notas da versão 6.13.0

As notas da versão a seguir incluem informações para a versão 6.13.0 do Amazon EMR. As alterações são referentes à versão 6.12.0. Para obter informações sobre o cronograma da versão, consulte o Log de alterações 6.13.0.

Novos recursos

- Amazon EMR 6.13.0 supports Apache Spark 3.4.1, Apache Spark RAPIDS 23.06.0-amzn-1, CUDA Toolkit 11.8.0, and JupyterHub 1.5.0.

Alterações, melhorias e problemas resolvidos

- A versão 6.13.0 aprimora o daemon de gerenciamento de logs do Amazon EMR para garantir que todos os logs sejam carregados regularmente para o Amazon S3 quando um comando de encerramento de cluster é emitido. Isso facilita o encerramento mais rápido do cluster.
- A versão 6.13.0 aprimora os recursos de gerenciamento de logs do Amazon EMR para garantir o carregamento consistente e oportuno de todos os arquivos de log para o Amazon S3. Isso beneficia especialmente os clusters do EMR de execução prolongada.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací fico (Seul), Ásia-Pacífico

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Osaka), Europa (Londres) , Ásia-Pacífico (Melbourn e), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (Oregon) AWS GovCloud , Oeste dos EUA (Norte da Califórnia), Ásia-Pací fico (AWS GovCloud Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv), Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha) , Europa (Estocolm o), Europa (Zurique) , Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Ásia-Pací fico (Jacarta) Ásia-Pacífico (Melbourne), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central), Oeste do Canadá (Calgary) AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud , (Leste dos EUA) AWS GovCloud , (Oeste dos EUA), China (Pequim), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Centro Leste (Emirados Árabes Unidos), Ásia-Pacífico (Melbourne),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Israel (Tel Aviv), Oeste do Canadá (Calgary), Europa (Zurique)
2.0.2024 001.0	4.14.352	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 816.0	4.14.350	21 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 809.0	4.14.349	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 719.0	4.14.348	25 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 223.0	4.14.336	8 de março de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 131.0	4.14.336	14 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 124.0	4.14.336	7 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 109.0	4.14.334	24 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 218.0	4.14.330	2 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 206.0	4.14.330	22 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 116.0	4.14.328	11 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 101.0	4.14.327	16 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 020.1	4.14.326	7 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 012.1	4.14.326	26 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 926.0	4.14.322	19 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 906.0	4.14.322	4 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 808.0	4.14.320	24 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

Versões padrão do Java 6.13.0

As versões 6.12.0 e superiores do Amazon EMR são compatíveis com todas as aplicações que têm o Amazon Corretto 8 por padrão, exceto o Trino. Para Trino, o Amazon EMR é compatível com o Amazon Corretto 17 por padrão, desde a versão 6.9.0 do Amazon EMR. O Amazon EMR também é compatível com algumas aplicações com o Amazon Corretto 11 e 17. Essas aplicações estão listadas na tabela a seguir. Se você quiser alterar a JVM padrão no cluster, siga as instruções em Configurar aplicações para usar uma Máquina Virtual Java específica para cada aplicação executada no cluster. Você só pode usar uma versão de runtime do Java para um cluster. O Amazon EMR não é compatível com a execução de diferentes nós ou aplicações em diferentes versões de runtime no mesmo cluster.

Embora o Amazon EMR seja compatível com o Amazon Corretto 11 e 17 no Apache Spark, no Apache Hadoop e no Apache Hive, a performance de algumas workloads pode diminuir quando você usa estas versões do Corretto. Recomendamos que você teste suas workloads antes de alterar os padrões.

A seguinte tabela mostra as versões padrão do Java para aplicações no Amazon EMR 6.13.0:

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
Delta	17, 11, 8
Flink	11, 8
Ganglia	8
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11, 8

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez	17, 11, 8
Trino	17
Zeppelin	8
Pig	8
Zookeeper	8

Versões do componente 6.13.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado

`myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.2	SDK do Amazon SageMaker Spark
delta	2.4.0	O Delta Lake é um formato de tabela aberto para grandes conjuntos de dados analíticos
delta-standalone-connectors	0.6.0	Os conectores do Delta oferecem runtimes diferente s para integrar o Delta Lake a mecanismos, como Flink, Hive e Presto.
emr-ddb	5.1.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.6.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.10.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.7.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.27.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	2.6.0	EMR S3Select Connector

Componente	Versão	Descrição
emr-wal-cli	1.1.0	Cli usada para lista/exclusão de emrwal.
emrfs	2.58.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.17.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.17.0	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	3.3.3-amzn-5	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.

Componente	Versão	Descrição
hadoop-hdfs-datanode	3.3.3-amzn-5	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.3.3-amzn-5	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.3.3-amzn-5	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	3.3.3-amzn-5	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.3.3-amzn-5	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.3.3-amzn-5	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	3.3.3-amzn-5	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.3.3-amzn-5	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	3.3.3-amzn-5	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.

Componente	Versão	Descrição
hadoop-yarn-timeline-server	3.3.3-amzn-5	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	2.4.17-amzn-1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	2.4.17-amzn-1	Serviço para atender a uma ou mais HBase regiões.
hbase-client	2.4.17-amzn-1	HBase cliente de linha de comando.
hbase-rest-server	2.4.17-amzn-1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.4.17-amzn-1	Serviço que fornece um endpoint Thrift para. HBase
hbase-operator-tools	2.4.17-amzn-1	Ferramenta de reparo para HBase clusters Apache.
hcatalog-client	3.1.3-amzn-6	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.3-amzn-6	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	3.1.3-amzn-6	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.3-amzn-6	O cliente da linha de comando do Hive.
hive-hbase	3.1.3-amzn-6	Cliente do Hive-hbase.
hive-metastore-server	3.1.3-amzn-6	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.3-amzn-6	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.13.1-amzn-1	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.13.1-amzn-1	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-trino	0.13.1-amzn-1	Biblioteca de pacotes para execução do Presto com o Hudi.
hudi-spark	0.13.1-amzn-1	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.11.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop

Componente	Versão	Descrição
iceberg	1.3.0-amzn-1	Apache Iceberg é um formato de tabela aberta para conjuntos de dados analíticos imensos
jupyterhub	1.5.0	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.1-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mxnet	1.9.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.
nvidia-cuda	11.8.0	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.7.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	5.1.3	As bibliotecas do phoenix para servidor e cliente

Componente	Versão	Descrição
phoenix-connectors	5.1.3	Conectores do Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.281-amzn-1	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.281-amzn-1	O serviço que executa partes de uma consulta.
presto-client	0.281-amzn-1	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
trino-coordinator	414-amzn-1	Serviço que aceita consultas e gerencia a execução de consultas entre trino-workers.
trino-worker	414-amzn-1	O serviço que executa partes de uma consulta.

Componente	Versão	Descrição
trino-client	414-amzn-1	Cliente de linha de comando do Trino instalado em mestres de espera de um cluster de alta disponibilidade (HA) em que o servidor do Trino não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.0.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	3.4.1-amzn-0	Os clientes da linha de comando do Spark.
spark-history-server	3.4.1-amzn-0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.4.1-amzn-0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.4.1-amzn-0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
spark-rapids	23.06.0-amzn-1	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs

Componente	Versão	Descrição
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.11.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.10.2-amzn-4	O aplicativo e as bibliotecas do YARN para Tez.
tez-on-worker	0.10.2-amzn-4	A aplicação e as bibliotecas do YARN nós de processam ento.
webserver	2.4.41+	O servidor Apache HTTP.
zeppelin-server	0.10.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.5.10	O serviço centralizado de manutenção de informações de configuração, nomenclat ura, fornecimento de sincroniz ação distribuída, e fornecime nto de serviços de grupo.
zookeeper-client	3.5.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 6.13.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-6.13.0

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.

Classificações	Descrição	Ações de reconfiguração
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.

Classificações	Descrição	Ações de reconfiguração
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.

Classificações	Descrição	Ações de reconfiguração
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classificações	Descrição	Ações de reconfiguração
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.

Classificações	Descrição	Ações de reconfiguração
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Alterar a página inicial do KMS do Hadoop	Not available.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.	Not available.
iceberg-defaults	Alterar os valores no arquivo iceberg-defaults.conf do Iceberg.	Not available.
delta-defaults	Alterar os valores no arquivo delta-defaults.conf do Delta.	Not available.

Classificações	Descrição	Ações de reconfiguração
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j2	Alterar as configurações de log4j2.properties do Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.

Classificações	Descrição	Ações de reconfiguração
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-lakeformation	Alterar os valores no arquivo lakeformation.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
trino-log	Alterar os valores no arquivo log.properties do Trino.	Restarts Trino-Server (for Trino)
trino-config	Alterar os valores no arquivo config.properties do Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Alterar os valores no arquivo password-authenticator.properties do Trino.	Restarts Trino-Server (for Trino)
trino-env	Alterar os valores no arquivo trino-env.sh do Trino.	Restarts Trino-Server (for Trino)

Classificações	Descrição	Ações de reconfiguração
trino-node	Alterar os valores no arquivo node.properties do Trino.	Not available.
trino-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Trino.	Not available.
trino-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Trino.	Not available.
trino-connector-delta	Alterar os valores no arquivo delta.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Alterar os valores no arquivo hive.properties do Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Alterar os valores no arquivo exchange-manager.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Alterar os valores no arquivo iceberg.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Alterar os valores no arquivo hudi.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Alterar os valores no arquivo jmx.properties do Trino.	Not available.
trino-connector-kafka	Alterar os valores no arquivo kafka.properties do Trino.	Not available.
trino-connector-localfile	Alterar os valores no arquivo localfile.properties do Trino.	Not available.
trino-connector-memory	Alterar os valores no arquivo memory.properties do Trino.	Not available.

Classificações	Descrição	Ações de reconfiguração
trino-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Trino.	Not available.
trino-connector-mysql	Alterar os valores no arquivo mysql.properties do Trino.	Not available.
trino-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Trino.	Not available.
trino-connector-raptor	Alterar os valores no arquivo raptor.properties do Trino.	Not available.
trino-connector-redis	Alterar os valores no arquivo redis.properties do Trino.	Not available.
trino-connector-redshift	Alterar os valores no arquivo redshift.properties do Trino.	Not available.
trino-connector-tpch	Alterar os valores no arquivo tpch.properties do Trino.	Not available.
trino-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Trino.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Alterar os valores no arquivo kms-logback.xml do Ranger KMS.	Not available.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
spark	As configurações administr adas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j2	Alterar os valores no arquivo log4j2.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie and HiveServe r2.

Classificações	Descrição	Ações de reconfiguração
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Log de alterações 6.13.0

Log de alterações da versão 6.13.0 e das notas da versão

Data	Event	Descrição
2023-09-23	Implantação concluída	Amazon EMR 6.13.0 totalmente implantado em todas as regiões compatíveis
2023-09-12	Publicação de documentos	Notas da versão 6.13.0 do Amazon EMR publicadas pela primeira vez
2023-09-01	Lançamento inicial	A versão 6.13.0 do Amazon EMR foi implantada pela primeira vez em regiões comerciais iniciais

Versão 6.12.0 do Amazon EMR

Versões da aplicação 6.12.0

Essa versão inclui os seguintes aplicativos: Delta, Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, SqoopTensorFlow, Tez, TrinoZeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-6.12.0	emr-6.11.1	emr-6.11.0	emr-6.10.1
AWS SDK for Java	1.12.490	1.12.446	1.12.446	1.12.397
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.4.0	2.2.0	2.2.0	2.2.0
Flink	1.17.0	1.16.0	1.16.0	1.16.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.17	2.4.15	2.4.15	2.4.15
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.13.1-amzn-0	0.13.0-amzn-0	0.13.0-amzn-0	0.12.2-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.10.0
Iceberg	1.3.0-amzn-0	1.2.0-amzn-0	1.2.0-amzn-0	1.1.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.5.0
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1

	emr-6.12.0	emr-6.11.1	emr-6.11.0	emr-6.10.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,281	0,279	0,279	0,278
Spark	3.4.0	3.3.2	3.3.2	3.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	414	410	410	403
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

Notas da versão 6.12.0

As notas da versão a seguir incluem informações para a versão 6.12.0 do Amazon EMR. As alterações são referentes à versão 6.11.0. Para obter informações sobre o cronograma da versão, consulte o Log de alterações 6.12.0.

Novos recursos

- Amazon EMR 6.12.0 supports Apache Spark 3.4.0, Apache Spark RAPIDS 23.06.0-amzn-0, CUDA 11.8.0, Apache Hudi 0.13.1-amzn-0, Apache Iceberg 1.3.0-amzn-0, Trino 414, and PrestoDB 0.281.
- As versões 6.12.0 e superiores do Amazon EMR oferecem suporte à integração LDAP com Apache Livy, Apache Hive through HiveServer 2 (), Trino, Presto e Hue. HS2 Você também pode

instalar o Apache Spark e o Apache Hadoop em um cluster do EMR que use a versão 6.12.0 ou superior e configurá-los para usar LDAP. Para obter mais informações, consulte Usar servidores do Active Directory ou LDAP para autenticação com o Amazon EMR.

Alterações, melhorias e problemas resolvidos

- As versões 6.12.0 e superiores do Amazon EMR oferecem suporte ao runtime do Java 11 para o Flink. Para obter mais informações, consulte Configurar o Flink para ser executado com o Java 11.
- A versão 6.12.0 adiciona um mecanismo de nova tentativa ao fluxo de trabalho de escalabilidade de clusters para os clusters do EMR que executam Presto ou Trino. Essa melhoria reduz o risco de que o redimensionamento do cluster fique paralisado indefinidamente devido a uma única falha na operação de redimensionamento. Ela também aprimora a utilização dos clusters, porque seu cluster aumenta e reduz a escala verticalmente com mais rapidez.
- A versão 6.12.0 corrige um problema em que as operações de redução da escala verticalmente do cluster podem ficar paralisadas quando um nó central que está passando por uma desativação tranquila se torna não íntegro por qualquer motivo antes de ser totalmente desativado.
- A versão 6.12.0 melhora a lógica de redução da escala verticalmente do cluster para que o cluster não tente reduzir a escala verticalmente dos nós centrais abaixo da configuração do fator de replicação do HDFS para o cluster. Isso se alinha aos seus requisitos de redundância de dados e reduz a probabilidade de uma operação de escalabilidade paralisar.
- A versão 6.12.0 aprimora a performance e a eficiência do serviço de monitoramento de integridade do Amazon EMR, ao aumentar a velocidade com que ele registra em log as mudanças de estado das instâncias. Essa melhoria reduz a probabilidade de degradação do desempenho dos nós do cluster que estão executando várias ferramentas de cliente ou aplicações de terceiros personalizadas.
- A versão 6.12.0 melhora a performance do daemon de gerenciamento de logs no cluster para o Amazon EMR. Como resultado, existe uma menor probabilidade de degradação da performance com clusters do EMR que executam etapas com alta simultaneidade.
- Com a versão 6.12.0 do Amazon EMR, o daemon de gerenciamento de logs foi atualizado para identificar todos os logs que estão em uso ativo com identificadores de arquivos abertos no armazenamento da instância local e nos processos associados. Essa atualização garante que o Amazon EMR exclua adequadamente os arquivos e recupere o espaço de armazenamento depois que os logs são arquivados no Amazon S3.
- A versão 6.12.0 inclui um aprimoramento do daemon de gerenciamento de logs que exclui diretórios de etapas vazios e não utilizados no sistema de arquivos de cluster local. Um número

excessivamente grande de diretórios vazios pode degradar a performance dos daemons do Amazon EMR e resultar na utilização excessiva do disco.

- A versão 6.12.0 permite a alternância de logs do servidor de linha do tempo do YARN. Isso minimiza os cenários de utilização excessiva do disco, especialmente para clusters de execução prolongada.
- O tamanho padrão do volume raiz aumentou para 15 GB nas versões 6.10.0 e superiores do Amazon EMR. O tamanho padrão do volume raiz das versões anteriores é de 10 GB.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Europa (Londres)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			, Ásia-Pacífico (Melbourn e), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (Oregon) AWS GovCloud , Oeste dos EUA (Norte da Califórnia), Ásia-Pací fico (AWS GovCloud Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv), Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha) , Europa (Estocolm o), Europa (Zurique) , Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Ásia-Pací fico (Jacarta) Ásia-Pacífico (Melbourne), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central), Oeste do Canadá (Calgary) AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud , (Leste dos EUA) AWS GovCloud , (Oeste dos EUA), China (Pequim), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Centro Leste (Emirados Árabes Unidos), Ásia-Pacífico (Melbourne),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Israel (Tel Aviv), Oeste do Canadá (Calgary), Europa (Zurique)
2.0.2024 001.0	4.14.352	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 816.0	4.14.350	21 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 809.0	4.14.349	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 719.0	4.14.348	25 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 223.0	4.14.336	8 de março de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 131.0	4.14.336	14 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 124.0	4.14.336	7 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 109.0	4.14.334	24 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 218.0	4.14.330	2 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 206.0	4.14.330	22 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 116.0	4.14.328	11 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 101.0	4.14.327	16 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 020.1	4.14.326	7 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 012.1	4.14.326	26 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 926.0	4.14.322	19 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 906.0	4.14.322	4 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 822.0	4.14.322	30 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 808.0	4.14.320	24 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 727.0	4.14.320	14 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 719.0	4.14.320	2 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 628.0	4.14.318	12 de julho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta) , África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central)

Versões padrão do Java 6.12.0

As versões 6.12.0 e superiores do Amazon EMR são compatíveis com todas as aplicações que têm o Amazon Corretto 8 por padrão, exceto o Trino. Para Trino, o Amazon EMR é compatível com o Amazon Corretto 17 por padrão, desde a versão 6.9.0 do Amazon EMR. O Amazon EMR também é compatível com algumas aplicações com o Amazon Corretto 11 e 17. Essas aplicações estão listadas na tabela a seguir. Se você quiser alterar a JVM padrão no cluster, siga as instruções em Configurar aplicações para usar uma Máquina Virtual Java específica para cada aplicação executada no cluster. Você só pode usar uma versão de runtime do Java para um cluster. O Amazon EMR não é compatível com a execução de diferentes nós ou aplicações em diferentes versões de runtime no mesmo cluster.

Embora o Amazon EMR seja compatível com o Amazon Corretto 11 e 17 no Apache Spark, no Apache Hadoop e no Apache Hive, a performance de algumas workloads pode diminuir quando você usa estas versões do Corretto. Recomendamos que você teste suas workloads antes de alterar os padrões.

A seguinte tabela mostra as versões padrão do Java para aplicações no Amazon EMR 6.12.0:

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
Delta	17, 11, 8
Flink	11, 8
Ganglia	8
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11, 8

Aplicação	Versão Java/Amazon Corretto (o padrão é negrito)
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez	17, 11, 8
Trino	17
Zeppelin	8
Pig	8
Zookeeper	8

Versões de componente 6.12.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado

`myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.2	SDK do Amazon SageMaker Spark
delta	2.4.0	O Delta Lake é um formato de tabela aberto para grandes conjuntos de dados analíticos
delta-standalone-connectors	0.6.0	Os conectores do Delta oferecem runtimes diferente s para integrar o Delta Lake a mecanismos, como Flink, Hive e Presto.
emr-ddb	5.1.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.5.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.9.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.7.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.26.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	2.5.0	EMR S3Select Connector

Componente	Versão	Descrição
emr-wal-cli	1.1.0	Cli usada para lista/exclusão de emrwal.
emrfs	2.57.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.17.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.17.0	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	3.3.3-amzn-4	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.

Componente	Versão	Descrição
hadoop-hdfs-datanode	3.3.3-amzn-4	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.3.3-amzn-4	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.3.3-amzn-4	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	3.3.3-amzn-4	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.3.3-amzn-4	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.3.3-amzn-4	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	3.3.3-amzn-4	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.3.3-amzn-4	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	3.3.3-amzn-4	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.

Componente	Versão	Descrição
hadoop-yarn-timeline-server	3.3.3-amzn-4	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	2.4.17-amzn-0	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	2.4.17-amzn-0	Serviço para atender a uma ou mais HBase regiões.
hbase-client	2.4.17-amzn-0	HBase cliente de linha de comando.
hbase-rest-server	2.4.17-amzn-0	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.4.17-amzn-0	Serviço que fornece um endpoint Thrift para. HBase
hbase-operator-tools	2.4.17-amzn-0	Ferramenta de reparo para HBase clusters Apache.
hcatalog-client	3.1.3-amzn-5	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.3-amzn-5	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	3.1.3-amzn-5	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.3-amzn-5	O cliente da linha de comando do Hive.
hive-hbase	3.1.3-amzn-5	Cliente do Hive-hbase.
hive-metastore-server	3.1.3-amzn-5	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.3-amzn-5	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.13.1-amzn-0	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.13.1-amzn-0	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-trino	0.13.1-amzn-0	Biblioteca de pacotes para execução do Presto com o Hudi.
hudi-spark	0.13.1-amzn-0	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.11.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop

Componente	Versão	Descrição
iceberg	1.3.0-amzn-0	Apache Iceberg é um formato de tabela aberta para conjuntos de dados analíticos imensos
jupyterhub	1.4.1	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.1-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mxnet	1.9.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.
nvidia-cuda	11.8.0	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.7.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	5.1.3	As bibliotecas do phoenix para servidor e cliente

Componente	Versão	Descrição
phoenix-connectors	5.1.3	Conectores do Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.281-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.281-amzn-0	O serviço que executa partes de uma consulta.
presto-client	0.281-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
trino-coordinator	414-amzn-0	Serviço que aceita consultas e gerencia a execução de consultas entre trino-workers.
trino-worker	414-amzn-0	O serviço que executa partes de uma consulta.

Componente	Versão	Descrição
trino-client	414-amzn-0	Cliente de linha de comando do Trino instalado em mestres de espera de um cluster de alta disponibilidade (HA) em que o servidor do Trino não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.0.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	3.4.0-amzn-0	Os clientes da linha de comando do Spark.
spark-history-server	3.4.0-amzn-0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.4.0-amzn-0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.4.0-amzn-0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
spark-rapids	23.06.0-amzn-0	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs

Componente	Versão	Descrição
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.11.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.10.2-amzn-3	O aplicativo e as bibliotecas do YARN para Tez.
tez-on-worker	0.10.2-amzn-3	A aplicação e as bibliotecas do YARN nós de processam ento.
webserver	2.4.41+	O servidor Apache HTTP.
zeppelin-server	0.10.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.5.10	O serviço centralizado de manutenção de informações de configuração, nomenclat ura, fornecimento de sincroniz ação distribuída, e fornecime nto de serviços de grupo.
zookeeper-client	3.5.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 6.12.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-6.12.0

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.

Classificações	Descrição	Ações de reconfiguração
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.

Classificações	Descrição	Ações de reconfiguração
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.

Classificações	Descrição	Ações de reconfiguração
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classificações	Descrição	Ações de reconfiguração
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.

Classificações	Descrição	Ações de reconfiguração
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Alterar a página inicial do KMS do Hadoop	Not available.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.	Not available.
iceberg-defaults	Alterar os valores no arquivo iceberg-defaults.conf do Iceberg.	Not available.
delta-defaults	Alterar os valores no arquivo delta-defaults.conf do Delta.	Not available.

Classificações	Descrição	Ações de reconfiguração
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j2	Alterar as configurações de log4j2.properties do Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.

Classificações	Descrição	Ações de reconfiguração
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-lakeformation	Alterar os valores no arquivo lakeformation.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
trino-log	Alterar os valores no arquivo log.properties do Trino.	Restarts Trino-Server (for Trino)
trino-config	Alterar os valores no arquivo config.properties do Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Alterar os valores no arquivo password-authenticator.properties do Trino.	Restarts Trino-Server (for Trino)
trino-env	Alterar os valores no arquivo trino-env.sh do Trino.	Restarts Trino-Server (for Trino)

Classificações	Descrição	Ações de reconfiguração
trino-node	Alterar os valores no arquivo node.properties do Trino.	Not available.
trino-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Trino.	Not available.
trino-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Trino.	Not available.
trino-connector-delta	Alterar os valores no arquivo delta.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Alterar os valores no arquivo hive.properties do Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Alterar os valores no arquivo exchange-manager.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Alterar os valores no arquivo iceberg.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Alterar os valores no arquivo hudi.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Alterar os valores no arquivo jmx.properties do Trino.	Not available.
trino-connector-kafka	Alterar os valores no arquivo kafka.properties do Trino.	Not available.
trino-connector-localfile	Alterar os valores no arquivo localfile.properties do Trino.	Not available.
trino-connector-memory	Alterar os valores no arquivo memory.properties do Trino.	Not available.

Classificações	Descrição	Ações de reconfiguração
trino-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Trino.	Not available.
trino-connector-mysql	Alterar os valores no arquivo mysql.properties do Trino.	Not available.
trino-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Trino.	Not available.
trino-connector-raptor	Alterar os valores no arquivo raptor.properties do Trino.	Not available.
trino-connector-redis	Alterar os valores no arquivo redis.properties do Trino.	Not available.
trino-connector-redshift	Alterar os valores no arquivo redshift.properties do Trino.	Not available.
trino-connector-tpch	Alterar os valores no arquivo tpch.properties do Trino.	Not available.
trino-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Trino.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Alterar os valores no arquivo kms-logback.xml do Ranger KMS.	Not available.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
spark	As configurações administr adas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j2	Alterar os valores no arquivo log4j2.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie and HiveServe r2.

Classificações	Descrição	Ações de reconfiguração
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Log de alterações 6.12.0

Log de alterações da versão 6.12.0 e das notas da versão

Data	Event	Descrição
2023-07-27	Atualizar a documentação	Atualizar as opções Java para 6.12 e adicionar o tutorial do Oozie para atualizar a JVM
2023-07-21	Implantação concluída	Amazon EMR 6.12.0 totalment e implantado em todas as regiões compatíveis
2023-07-21	Publicação de documentos	Notas da versão 6.12.0 do Amazon EMR publicadas pela primeira vez
2023-07-12	Lançamento inicial	A versão 6.12.0 do Amazon EMR foi implantada pela primeira vez em regiões comerciais iniciais

Versão 6.11.1 do Amazon EMR

Versões 6.11.1 da aplicação

Essa versão inclui os seguintes aplicativos: Delta, Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, SqoopTensorFlow, Tez, TrinoZeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR

- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-6.11.1	emr-6.11.0	emr-6.10.1	emr-6.10.0
AWS SDK for Java	1.12.446	1.12.446	1.12.397	1.12.397
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.2.0	2.2.0	2.2.0	2.2.0
Flink	1.16.0	1.16.0	1.16.0	1.16.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.15	2.4.15	2.4.15	2.4.15
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.13.0-amzn-0	0.13.0-amzn-0	0.12.2-amzn-0	0.12.2-amzn-0
Hue	4.11.0	4.11.0	4.10.0	4.10.0
Iceberg	1.2.0-amzn-0	1.2.0-amzn-0	1.1.0-amzn-0	1.1.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.4.1	1.4.1	1.5.0	1.5.0

	emr-6.11.1	emr-6.11.0	emr-6.10.1	emr-6.10.0
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,279	0,279	0,278	0,278
Spark	3.3.2	3.3.2	3.3.1	3.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	410	410	403	403
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

Notas da versão 6.11.1

As notas da versão a seguir incluem informações para a versão 6.11.1 do Amazon EMR. As alterações são referentes à versão 6.11.0. Para obter informações sobre o cronograma da versão, consulte o Log de alterações 6.11.1.

Alterações, melhorias e problemas resolvidos

- Devido à contenção de bloqueio, um nó pode entrar em um deadlock se for adicionado ou removido ao mesmo tempo em que tenta ser desativado. Como resultado, o Hadoop Resource

Manager (YARN) deixa de responder e afeta todos os contêineres de entrada e em execução no momento.

- Esta versão inclui uma alteração que permite que clusters de alta disponibilidade se recuperem de um estado de falha após a reinicialização.
- Esta versão inclui correções de segurança para Hue e. HBase
- Esta versão corrige um problema em que clusters que estão executando workloads no Spark com o Amazon EMR podem receber silenciosamente resultados incorretos com `contains`, `startsWith`, `endsWith` e `like`. Esse problema ocorre quando você usa as expressões em campos particionados que têm metadados no Hive3 Metastore Server (HMS) no Amazon EMR.
- Esta versão corrige um problema com controle de utilização no lado do Glue quando não há funções definidas pelo usuário (UDF).
- Esta versão corrige um problema que exclui logs de contêineres pelo serviço de agregação de logs de nó antes que o pusher de logs possa enviá-los para o S3 em caso de desativação do YARN.
- Esta versão corrige um problema com as métricas do FairShare Scheduler quando o Node Label está habilitado para o Hadoop.
- Esta versão corrige um problema que afetou a performance do Spark quando você definiu um valor de `true` padrão para a configuração `spark.yarn.heterogeneousExecutors.enabled` no `spark-defaults.conf`.
- Esta versão corrige um problema com a falha do Reduce Task em ler dados embaralhados. O problema causou falhas na consulta do Hive com um erro de memória corrompida.
- Esta versão adiciona um mecanismo de nova tentativa ao fluxo de trabalho de escalabilidade de clusters para clusters do EMR que executam o Presto ou o Trino. Essa melhoria reduz o risco de que o redimensionamento do cluster fique paralisado indefinidamente devido a uma única falha na operação de redimensionamento. Ela também aprimora a utilização dos clusters, porque seu cluster aumenta e reduz a escala verticalmente com mais rapidez.
- Esta versão melhora a lógica de redução da escala verticalmente do cluster para que o cluster não tente reduzir a escala verticalmente dos nós centrais abaixo da configuração do fator de replicação do HDFS para o cluster. Isso se alinha aos seus requisitos de redundância de dados e reduz a probabilidade de uma operação de escalabilidade paralisar.
- O daemon de gerenciamento de logs foi atualizado para identificar todos os logs que estão em uso ativo com identificadores de arquivos abertos no armazenamento da instância local e nos processos associados. Essa atualização garante que o Amazon EMR exclua adequadamente os arquivos e recupere o espaço de armazenamento depois que os logs são arquivados no Amazon S3.

- Esta versão inclui um aprimoramento do daemon de gerenciamento de logs que exclui diretórios de etapas vazios e não utilizados no sistema de arquivos de cluster local. Um número excessivamente grande de diretórios vazios pode degradar a performance dos daemons do Amazon EMR e resultar na utilização excessiva do disco.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Malásia), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Malásia), Europa (Londres), Ásia-Pacífico

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Melbourne), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (AWS GovCloud AWS GovCloud Oregon), México (Central) , Oeste dos EUA (Norte da Califórnia), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv), Ásia-Pacífico (Taipei), Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Hyderabad), Ásia-Pacífico (Jacarta) Arta), Ásia-Pacífico (Malásia), Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Malásia) , Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (EAU), Canadá (Central) , Oeste do Canadá

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Calgary), (Leste dos EUA), AWS GovCloud (Oeste dos EUA), China (AWS GovCloud Pequim), China (Ningxia), Israel (Tel Aviv), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Centro Leste (Emirados Árabes Unidos), Ásia-Pacífico (Melbourne),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Israel (Tel Aviv), Oeste do Canadá (Calgary), Europa (Zurique), Ásia-Pacífico (Malásia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 001.0	4.14.352	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 816.0	4.14.350	21 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 809.0	4.14.349	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 719.0	4.14.348	25 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 223.0	4.14.336	8 de março de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 131.0	4.14.336	14 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 124.0	4.14.336	7 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 109.0	4.14.334	24 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 218.0	4.14.330	2 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 206.0	4.14.330	22 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 116.0	4.14.328	11 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 101.0	4.14.327	16 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 020.1	4.14.326	7 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 012.1	4.14.326	26 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 926.0	4.14.322	19 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 906.0	4.14.322	4 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 822.0	4.14.322	30 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 808.0	4.14.320	24 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 727.0	4.14.320	14 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central)

Versões de componente 6.11.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a

versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *CommunityVersion*-amzn-*EmrVersion*. O *EmrVersion* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.2	SDK do Amazon SageMaker Spark
delta	2.2.0	O Delta Lake é um formato de tabela aberto para grandes conjuntos de dados analíticos
delta-standalone-connectors	0.6.0	Os conectores do Delta oferecem runtimes diferente s para integrar o Delta Lake a mecanismos, como Flink, Hive e Presto.
emr-ddb	5.1.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.8.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.7.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway

Componente	Versão	Descrição
emr-s3-dist-cp	2.25.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	2.4.0	EMR S3Select Connector
emr-wal-cli	1.1.0	Cli usada para lista/exclusão de emrwal.
emrfs	2.56.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.16.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.16.0	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.

Componente	Versão	Descrição
hadoop-client	3.3.3-amzn-3.1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	3.3.3-amzn-3.1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.3.3-amzn-3.1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.3.3-amzn-3.1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	3.3.3-amzn-3.1	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.3.3-amzn-3.1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.3.3-amzn-3.1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	3.3.3-amzn-3.1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.3.3-amzn-3.1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	3.3.3-amzn-3.1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	3.3.3-amzn-3.1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	2.4.15-amzn-1.1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	2.4.15-amzn-1.1	Serviço para atender a uma ou mais HBase regiões.
hbase-client	2.4.15-amzn-1.1	HBase cliente de linha de comando.
hbase-rest-server	2.4.15-amzn-1.1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.4.15-amzn-1.1	Serviço que fornece um endpoint Thrift para. HBase
hbase-operator-tools	2.4.15-amzn-1.1	Ferramenta de reparo para HBase clusters Apache.
hcatalog-client	3.1.3-amzn-4.1	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.

Componente	Versão	Descrição
hcatalog-server	3.1.3-amzn-4.1	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	3.1.3-amzn-4.1	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.3-amzn-4.1	O cliente da linha de comando do Hive.
hive-hbase	3.1.3-amzn-4.1	Cliente do Hive-hbase.
hive-metastore-server	3.1.3-amzn-4.1	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.3-amzn-4.1	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.13.0-amzn-0	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.13.0-amzn-0	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-trino	0.13.0-amzn-0	Biblioteca de pacotes para execução do Presto com o Hudi.

Componente	Versão	Descrição
hudi-spark	0.13.0-amzn-0	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.11.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
iceberg	1.2.0-amzn-0	Apache Iceberg é um formato de tabela aberta para conjuntos de dados analíticos imensos
jupyterhub	1.4.1	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.1-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mxnet	1.9.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.
nvidia-cuda	11.8.0	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.

Componente	Versão	Descrição
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.5.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	5.1.2	As bibliotecas do phoenix para servidor e cliente
phoenix-connectors	5.1.2	Conectores do Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.279-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.279-amzn-0	O serviço que executa partes de uma consulta.
presto-client	0.279-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
trino-coordinator	410-amzn-0	Serviço que aceita consultas e gerencia a execução de consultas entre trino-workers.

Componente	Versão	Descrição
trino-worker	410-amzn-0	O serviço que executa partes de uma consulta.
trino-client	410-amzn-0	Cliente de linha de comando do Trino instalado em mestres de espera de um cluster de alta disponibilidade (HA) em que o servidor do Trino não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.0.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	3.3.2-amzn-0.1	Os clientes da linha de comando do Spark.
spark-history-server	3.3.2-amzn-0.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.3.2-amzn-0.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.3.2-amzn-0.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.

Componente	Versão	Descrição
spark-rapids	23.02.0-amzn-0	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.11.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.10.2-amzn-2.1	O aplicativo e as bibliotecas do YARN para Tez.
tez-on-worker	0.10.2-amzn-2.1	A aplicação e as bibliotecas do YARN nós de processam ento.
webserver	2.4.41+	O servidor Apache HTTP.
zeppelin-server	0.10.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.5.10	O serviço centralizado de manutenção de informações de configuração, nomenclat ura, fornecimento de sincroniz ação distribuída, e fornecime nto de serviços de grupo.
zookeeper-client	3.5.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 6.11.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-6.11.1

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop

Classificações	Descrição	Ações de reconfiguração
		Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.

Classificações	Descrição	Ações de reconfiguração
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.

Classificações	Descrição	Ações de reconfiguração
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hbase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classificações	Descrição	Ações de reconfiguração
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.

Classificações	Descrição	Ações de reconfiguração
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.	Not available.
iceberg-defaults	Alterar os valores no arquivo iceberg-defaults.conf do Iceberg.	Not available.
delta-defaults	Alterar os valores no arquivo delta-defaults.conf do Delta.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.

Classificações	Descrição	Ações de reconfiguração
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j2	Alterar as configurações de log4j2.properties do Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.

Classificações	Descrição	Ações de reconfiguração
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)

Classificações	Descrição	Ações de reconfiguração
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-lakeformation	Alterar os valores no arquivo lakeformation.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
trino-log	Alterar os valores no arquivo log.properties do Trino.	Restarts Trino-Server (for Trino)
trino-config	Alterar os valores no arquivo config.properties do Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Alterar os valores no arquivo password-authenticator.properties do Trino.	Restarts Trino-Server (for Trino)
trino-env	Alterar os valores no arquivo trino-env.sh do Trino.	Restarts Trino-Server (for Trino)
trino-node	Alterar os valores no arquivo node.properties do Trino.	Not available.

Classificações	Descrição	Ações de reconfiguração
trino-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Trino.	Not available.
trino-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Trino.	Not available.
trino-connector-delta	Alterar os valores no arquivo delta.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Alterar os valores no arquivo hive.properties do Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Alterar os valores no arquivo exchange-manager.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Alterar os valores no arquivo iceberg.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Alterar os valores no arquivo hudi.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Alterar os valores no arquivo jmx.properties do Trino.	Not available.
trino-connector-kafka	Alterar os valores no arquivo kafka.properties do Trino.	Not available.
trino-connector-localfile	Alterar os valores no arquivo localfile.properties do Trino.	Not available.
trino-connector-memory	Alterar os valores no arquivo memory.properties do Trino.	Not available.
trino-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Trino.	Not available.

Classificações	Descrição	Ações de reconfiguração
trino-connector-mysql	Alterar os valores no arquivo mysql.properties do Trino.	Not available.
trino-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Trino.	Not available.
trino-connector-raptor	Alterar os valores no arquivo raptor.properties do Trino.	Not available.
trino-connector-redis	Alterar os valores no arquivo redis.properties do Trino.	Not available.
trino-connector-redshift	Alterar os valores no arquivo redshift.properties do Trino.	Not available.
trino-connector-tpch	Alterar os valores no arquivo tpch.properties do Trino.	Not available.
trino-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Trino.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Alterar os valores no arquivo kms-logback.xml do Ranger KMS.	Not available.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
spark	As configurações administr adas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j2	Alterar os valores no arquivo log4j2.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie and HiveServe r2.

Classificações	Descrição	Ações de reconfiguração
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Log de alterações 6.11.1

Log de alterações da versão 6.11.1 e das notas da versão

Data	Event	Descrição
2023-08-30	Atualizar notas da versão	Adicionadas várias correções relacionadas ao ambiente de gerenciamento às notas da versão
2023-08-21	Publicação de documentos	Notas da versão 6.11.1 do Amazon EMR publicadas pela primeira vez
2023-08-16	Implantação concluída	Amazon EMR 6.11.1 totalment e implantado em todas as regiões compatíveis
2023-08-04	Lançamento inicial	A versão 6.11.1 do Amazon EMR foi implantada pela primeira vez em regiões comerciais limitadas

Versão 6.11.0 do Amazon EMR

Versões da aplicação 6.11.0

Essa versão inclui os seguintes aplicativos: Delta, Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, SqoopTensorFlow, Tez, TrinoZeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR

- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-6.11.0	emr-6.10.1	emr-6.10.0	emr-6.9.1
AWS SDK for Java	1.12.446	1.12.397	1.12.397	1.12.170
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.2.0	2.2.0	2.2.0	2.1.0
Flink	1.16.0	1.16.0	1.16.0	1.15.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.15	2.4.15	2.4.15	2.4.13
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.13.0-amzn-0	0.12.2-amzn-0	0.12.2-amzn-0	0.12.1-amzn-0
Hue	4.11.0	4.10.0	4.10.0	4.10.0
Iceberg	1.2.0-amzn-0	1.1.0-amzn-0	1.1.0-amzn-0	0.14.1-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0

	emr-6.11.0	emr-6.10.1	emr-6.10.0	emr-6.9.1
JupyterHub	1.4.1	1.5.0	1.5.0	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,279	0,278	0,278	0,276
Spark	3.3.2	3.3.1	3.3.1	3.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.10.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	410	403	403	398
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

Notas da versão 6.11.0

As notas da versão a seguir incluem informações para a versão 6.11.0 do Amazon EMR. As alterações são referentes à versão 6.10.0. Para obter informações sobre o cronograma da versão, consulte o log de alterações.

Novos recursos

- O Amazon EMR 6.11.0 é compatível com Apache Spark 3.3.2-amzn-0, Apache Spark RAPIDS 23.02.0-amzn-0, CUDA 11.8.0, Apache Hudi 0.13.0-amzn-0, Apache Iceberg 1.2.0-amzn-0, Trino 410-amzn-0 e PrestoDB 0.279-amzn-0.

Alterações, melhorias e problemas resolvidos

- Com o Amazon EMR 6.11.0, o conector do DynamoDB foi atualizado para a versão 5.0.0. A versão 5.0.0 usa AWS SDK for Java 2.x. As versões anteriores usavam AWS SDK para Java 1.x. Como resultado dessa atualização, recomendamos firmemente que você teste seu código antes de usar o conector do DynamoDB com o Amazon EMR 6.11.
- Quando o conector do DynamoDB para Amazon EMR 6.11.0 chama o serviço do DynamoDB, ele usa o valor da região que você fornece para a propriedade `dynamodb.endpoint`. Recomendamos que você também configure `dynamodb.region` quando usar `dynamodb.endpoint` e que ambas as propriedades tenham como destino a mesma Região da AWS. Se você usar `dynamodb.endpoint` e não configurar`dynamodb.region`, o conector do DynamoDB para Amazon EMR 6.11.0 retornará uma exceção de região inválida e tentará reconciliar suas Região da AWS informações do serviço de metadados de instância da Amazon (IMDS). EC2 Se o conector não conseguir recuperar a região do IMDS, o padrão será Leste dos EUA (Norte da Virgínia) (`us-east-1`). O erro a seguir é um exemplo da exceção de região inválida que você pode obter se não configurar adequadamente a `dynamodb.region` propriedade: `error software.amazon.awssdk.services.dynamodb.model.DynamoDbException: Credential should be scoped to a valid region.` Para obter mais informações sobre as classes afetadas pela AWS SDK para Java atualização para 2.x, consulte o commit Upgrade AWS SDK para Java from 1.x to 2.x (#175) no GitHub repositório do conector Amazon EMR - DynamoDB.
- Esta versão corrige um problema em que os dados da coluna se tornam `NULL` quando você usa o Delta Lake para armazenar dados da tabela Delta no Amazon S3 após a operação de renomeação da coluna. Para obter mais informações sobre esse atributo experimental no Delta Lake, consulte Operação de renomeação de coluna no Guia do usuário do Delta Lake.
- A versão 6.11.0 corrige um problema que pode ocorrer quando você cria um nó de borda ao replicar um dos nós primários de um cluster com vários nós primários. O nó de borda replicado pode causar atrasos nas operações de redução da escala verticalmente ou resultar em alta utilização de memória nos nós primários. Para obter mais informações sobre como criar um nó de

borda para se comunicar com seu cluster EMR, consulte Edge Node Creator no `aws-samples` repositório em. GitHub

- A versão 6.11.0 melhora o processo de automação que o Amazon EMR usa para remontar volumes do Amazon EBS em uma instância após uma reinicialização.
- A versão 6.11.0 corrige um problema que resultou em lacunas intermitentes nas métricas do Hadoop que o Amazon EMR publica na Amazon. CloudWatch
- A versão 6.11.0 corrige um problema com clusters do EMR em que uma atualização no arquivo de configuração do YARN que contém a lista de exclusão de nós do cluster é interrompida devido à utilização excessiva do disco. A atualização incompleta impede futuras operações de redução da escala verticalmente do cluster. Esta versão garante que o cluster permaneça íntegro e que as operações de escalabilidade funcionem conforme esperado.
- O tamanho padrão do volume raiz aumentou para 15 GB nas versões 6.10.0 e superiores do Amazon EMR. O tamanho padrão do volume raiz das versões anteriores é de 10 GB.
- O Hadoop 3.3.3 introduziu uma alteração no YARN (YARN-9608) que mantém os nós em que os contêineres eram executados em um estado de desativação até que a aplicação seja concluída. Essa alteração garante que dados locais, como dados embaralhados, não sejam perdidos e que você não precise executar o trabalho novamente. Essa abordagem também pode levar à subutilização de recursos em clusters com ou sem o ajuste de escala gerenciado habilitado.

  Com as versões 6.11.0 e superiores do Amazon EMR, além das versões 6.8.1, 6.9.1 e 6.10.1, o valor de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` é definido como `false` em `yarn-site.xml` para resolver esse problema.

  Embora a correção resolva os problemas introduzidos pelo YARN-9608, ela pode fazer com que os trabalhos do Hive falhem devido à perda de dados embaralhados em clusters com ajuste de escala gerenciado habilitado. Reduzimos esse risco nesta versão também ao configurar `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data` para workloads do Hive. Essa configuração só está disponível com as versões 6.11.0 e superiores do Amazon EMR.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

> **Note**
>
> Essa versão não recebe mais atualizações automáticas da AMI, pois foi substituída por uma ou mais versões de patch. A versão de patch é indicada pelo número após o segundo ponto decimal (`6.8.1`). Para ver se você está usando a versão de patch mais recente, verifique as versões disponíveis no Guia de lançamento ou verifique o menu suspenso de versões do Amazon EMR quando criar um cluster no console ou use a ação de API `ListReleaseLabels` ou da CLI `list-release-labels`. Para obter atualizações sobre novas versões, assine o feed RSS na página Novidades.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Europa (Londres) , Ásia-Pacífico (Melbourn e), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (Oregon) AWS GovCloud , Oeste dos EUA (Norte da Califórnia), Ásia-Pací fico (AWS GovCloud Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha) , Europa (Estocolm o), Europa (Zurique) , Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Ásia-Pací fico (Jacarta) Ásia-Pacífico (Melbourne), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central), Oeste do Canadá (Calgary) AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud , (Leste dos EUA) AWS GovCloud , (Oeste dos EUA), China (Pequim), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Centro Leste (Emirados Árabes Unidos), Ásia-Pacífico (Melbourne),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Israel (Tel Aviv), Oeste do Canadá (Calgary), Europa (Zurique)
2.0.2024 001.0	4.14.352	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 816.0	4.14.350	21 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 809.0	4.14.349	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 719.0	4.14.348	25 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)
2.0.2023 808.0	4.14.320	24 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), Israel (Tel Aviv).

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 727.0	4.14.320	14 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 719.0	4.14.320	2 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 628.0	4.14.318	12 de julho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Zurique), Europa (Milão), Europa (Espanha) , Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 612.0	4.14.314	23 de junho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Zurique), Europa (Milão), Europa (Espanha) , Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 504.1	4.14.313	16 de maio de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta) , África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central)

Versões de componente 6.11.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.2	SDK do Amazon SageMaker Spark
delta	2.2.0	O Delta Lake é um formato de tabela aberto para grandes conjuntos de dados analíticos
delta-standalone-connectors	0.6.0	Os conectores do Delta oferecem runtimes diferente s para integrar o Delta Lake a mecanismos, como Flink, Hive e Presto.
emr-ddb	5.1.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.

Componente	Versão	Descrição
emr-kinesis	3.8.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.7.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.25.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	2.4.0	EMR S3Select Connector
emr-wal-cli	1.1.0	Cli usada para lista/exclusão de emrwal.
emrfs	2.56.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.16.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.16.0	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	3.3.3-amzn-3	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	3.3.3-amzn-3	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.3.3-amzn-3	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.3.3-amzn-3	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	3.3.3-amzn-3	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.3.3-amzn-3	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.3.3-amzn-3	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider

Componente	Versão	Descrição
hadoop-mapred	3.3.3-amzn-3	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.3.3-amzn-3	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	3.3.3-amzn-3	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hadoop-yarn-timeline-server	3.3.3-amzn-3	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	2.4.15-amzn-1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	2.4.15-amzn-1	Serviço para atender uma ou mais HBase regiões.
hbase-client	2.4.15-amzn-1	HBase cliente de linha de comando.
hbase-rest-server	2.4.15-amzn-1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.4.15-amzn-1	Serviço que fornece um endpoint Thrift para. HBase

Componente	Versão	Descrição
hbase-operator-tools	2.4.15-amzn-1	Ferramenta de reparo para HBase clusters Apache.
hcatalog-client	3.1.3-amzn-4	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.3-amzn-4	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	3.1.3-amzn-4	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.3-amzn-4	O cliente da linha de comando do Hive.
hive-hbase	3.1.3-amzn-4	Cliente do Hive-hbase.
hive-metastore-server	3.1.3-amzn-4	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.3-amzn-4	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.13.0-amzn-0	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.

Componente	Versão	Descrição
hudi-presto	0.13.0-amzn-0	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-trino	0.13.0-amzn-0	Biblioteca de pacotes para execução do Presto com o Hudi.
hudi-spark	0.13.0-amzn-0	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.11.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
iceberg	1.2.0-amzn-0	Apache Iceberg é um formato de tabela aberta para conjuntos de dados analíticos imensos
jupyterhub	1.4.1	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.1-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mxnet	1.9.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.

Componente	Versão	Descrição
nvidia-cuda	11.8.0	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitações de fluxo de trabalho do Oozie.
opencv	4.5.0	Biblioteca de Visão Computacional de Código Aberto.
phoenix-library	5.1.2	As bibliotecas do phoenix para servidor e cliente
phoenix-connectors	5.1.2	Conectores do Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.279-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.279-amzn-0	O serviço que executa partes de uma consulta.

Componente	Versão	Descrição
presto-client	0.279-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
trino-coordinator	410-amzn-0	Serviço que aceita consultas e gerencia a execução de consultas entre trino-workers.
trino-worker	410-amzn-0	O serviço que executa partes de uma consulta.
trino-client	410-amzn-0	Cliente de linha de comando do Trino instalado em mestres de espera de um cluster de alta disponibilidade (HA) em que o servidor do Trino não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.0.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	3.3.2-amzn-0	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	3.3.2-amzn-0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.3.2-amzn-0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.3.2-amzn-0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
spark-rapids	23.02.0-amzn-0	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.11.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.10.2-amzn-2	O aplicativo e as bibliotecas do YARN para Tez.
tez-on-worker	0.10.2-amzn-2	A aplicação e as bibliotecas do YARN nós de processam ento.
webserver	2.4.41+	O servidor Apache HTTP.
zeppelin-server	0.10.1	O notebook baseado na web que permite um data analytics interativo.

Componente	Versão	Descrição
zookeeper-server	3.5.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.5.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 6.11.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-6.11.0

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Classificações	Descrição	Ações de reconfiguração
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.

Classificações	Descrição	Ações de reconfiguração
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.

Classificações	Descrição	Ações de reconfiguração
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.	Not available.
iceberg-defaults	Alterar os valores no arquivo iceberg-defaults.conf do Iceberg.	Not available.
delta-defaults	Alterar os valores no arquivo delta-defaults.conf do Delta.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j2	Alterar as configurações de log4j2.properties do Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-lakeformation	Alterar os valores no arquivo lakeformation.properties do Presto.	Restarts Presto-Server (for PrestoDB)

Classificações	Descrição	Ações de reconfiguração
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
trino-log	Alterar os valores no arquivo log.properties do Trino.	Restarts Trino-Server (for Trino)

Classificações	Descrição	Ações de reconfiguração
trino-config	Alterar os valores no arquivo config.properties do Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Alterar os valores no arquivo password-authenticator.properties do Trino.	Restarts Trino-Server (for Trino)
trino-env	Alterar os valores no arquivo trino-env.sh do Trino.	Restarts Trino-Server (for Trino)
trino-node	Alterar os valores no arquivo node.properties do Trino.	Not available.
trino-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Trino.	Not available.
trino-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Trino.	Not available.
trino-connector-delta	Alterar os valores no arquivo delta.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Alterar os valores no arquivo hive.properties do Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Alterar os valores no arquivo exchange-manager.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Alterar os valores no arquivo iceberg.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Alterar os valores no arquivo hudi.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Alterar os valores no arquivo jmx.properties do Trino.	Not available.

Classificações	Descrição	Ações de reconfiguração
trino-connector-kafka	Alterar os valores no arquivo kafka.properties do Trino.	Not available.
trino-connector-localfile	Alterar os valores no arquivo localfile.properties do Trino.	Not available.
trino-connector-memory	Alterar os valores no arquivo memory.properties do Trino.	Not available.
trino-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Trino.	Not available.
trino-connector-mysql	Alterar os valores no arquivo mysql.properties do Trino.	Not available.
trino-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Trino.	Not available.
trino-connector-raptor	Alterar os valores no arquivo raptor.properties do Trino.	Not available.
trino-connector-redis	Alterar os valores no arquivo redis.properties do Trino.	Not available.
trino-connector-redshift	Alterar os valores no arquivo redshift.properties do Trino.	Not available.
trino-connector-tpch	Alterar os valores no arquivo tpch.properties do Trino.	Not available.
trino-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Trino.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Alterar os valores no arquivo kms-logback.xml do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j2	Alterar os valores no arquivo log4j2.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie and HiveServe r2.
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.

Classificações	Descrição	Ações de reconfiguração
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Log de alterações 6.11.0

Log de alterações da versão 6.11.0 e das notas da versão

Data	Event	Descrição
2023-08-21	Atualizar	Problema corrigido introduzido com o Hadoop 3.3.3.
2023-07-26	Atualizar	Novos rótulos de versão `2.0.20230612.0` e `2.0.20230628.0` do sistema operacional.
2023-06-09	Implantação concluída	Amazon EMR 6.11.0 totalmente implantado em todas as regiões compatíveis
2023-06-09	Publicação de documentos	Notas da versão 6.11.0 do Amazon EMR publicadas pela primeira vez
2023-06-08	Lançamento inicial	A versão 6.11.0 do Amazon EMR foi implantada pela primeira vez em regiões comerciais iniciais

Versão 6.10.1 do Amazon EMR

Versões 6.10.1 da aplicação

Essa versão inclui os seguintes aplicativos: Delta, Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, SqoopTensorFlow, Tez, TrinoZeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-6.10.1	emr-6.10.0	emr-6.9.1	emr-6.9.0
AWS SDK for Java	1.12.397	1.12.397	1.12.170	1.12.170
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.2.0	2.2.0	2.1.0	2.1.0
Flink	1.16.0	1.16.0	1.15.2	1.15.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.15	2.4.15	2.4.13	2.4.13

	emr-6.10.1	emr-6.10.0	emr-6.9.1	emr-6.9.0
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.12.2-amzn-0	0.12.2-amzn-0	0.12.1-amzn-0	0.12.1-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.10.0
Iceberg	1.1.0-amzn-0	1.1.0-amzn-0	0.14.1-amzn-0	0.14.1-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,278	0,278	0,276	0,276
Spark	3.3.1	3.3.1	3.3.0	3.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.10.0	2.10.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2

	emr-6.10.1	emr-6.10.0	emr-6.9.1	emr-6.9.0
Trino (PrestoSQL)	403	403	398	398
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

Notas da versão 6.10.1

As notas da versão a seguir incluem informações para a versão 6.10.1 do Amazon EMR. As alterações são referentes à versão 6.10.0. Para obter informações sobre o cronograma da versão, consulte o Log de alterações 6.10.1.

Alterações, melhorias e problemas resolvidos

- Devido à contenção de bloqueio, um nó pode entrar em um deadlock se for adicionado ou removido ao mesmo tempo em que tenta ser desativado. Como resultado, o Hadoop Resource Manager (YARN) deixa de responder e afeta todos os contêineres de entrada e em execução no momento.
- O Hadoop 3.3.3 introduziu uma alteração no YARN (YARN-9608) que mantém os nós em que os contêineres eram executados em um estado de desativação até que a aplicação seja concluída. Essa alteração garante que dados locais, como dados embaralhados, não sejam perdidos e que você não precise executar o trabalho novamente. Essa abordagem também pode levar à subutilização de recursos em clusters com ou sem o ajuste de escala gerenciado habilitado.

  Com as versões 6.11.0 e superiores do Amazon EMR, além das versões 6.8.1, 6.9.1 e 6.10.1, o valor de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` é definido como `false` em `yarn-site.xml` para resolver esse problema.

  Embora a correção resolva os problemas introduzidos pelo YARN-9608, ela pode fazer com que os trabalhos do Hive falhem devido à perda de dados embaralhados em clusters com ajuste de escala gerenciado habilitado. Reduzimos esse risco nesta versão também ao configurar `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data` para workloads do Hive. Essa configuração só está disponível com as versões 6.11.0 e superiores do Amazon EMR.

- O coletor de métricas não enviará métricas para o ambiente de gerenciamento após o failover do nó primário em clusters com a configuração de grupos de instâncias.
- Esta versão inclui uma alteração que permite que clusters de alta disponibilidade se recuperem de um estado de falha após a reinicialização.
- Esta versão inclui correções de segurança para Hue e. HBase
- Esta versão corrige um problema em que clusters que estão executando workloads no Spark com o Amazon EMR podem receber silenciosamente resultados incorretos com `contains`, `startsWith`, `endsWith` e `like`. Esse problema ocorre quando você usa as expressões em campos particionados que têm metadados no Hive3 Metastore Server (HMS) no Amazon EMR.
- Esta versão corrige um problema com controle de utilização no lado do Glue quando não há funções definidas pelo usuário (UDF).
- Esta versão corrige um problema que exclui logs de contêineres pelo serviço de agregação de logs de nó antes que o pusher de logs possa enviá-los para o S3 em caso de desativação do YARN.
- Esta versão corrige um problema com as métricas do FairShare Scheduler quando o Node Label está habilitado para o Hadoop.
- Esta versão corrige um problema que afetou a performance do Spark quando você definiu um valor de `true` padrão para a configuração `spark.yarn.heterogeneousExecutors.enabled` no `spark-defaults.conf`.
- Esta versão corrige um problema com a falha do Reduce Task em ler dados embaralhados. O problema causou falhas na consulta do Hive com um erro de memória corrompida.
- Esta versão adiciona um mecanismo de nova tentativa ao fluxo de trabalho de escalabilidade de clusters para clusters do EMR que executam o Presto ou o Trino. Essa melhoria reduz o risco de que o redimensionamento do cluster fique paralisado indefinidamente devido a uma única falha na operação de redimensionamento. Ela também aprimora a utilização dos clusters, porque seu cluster aumenta e reduz a escala verticalmente com mais rapidez.
- Esta versão melhora a lógica de redução da escala verticalmente do cluster para que o cluster não tente reduzir a escala verticalmente dos nós centrais abaixo da configuração do fator de replicação do HDFS para o cluster. Isso se alinha aos seus requisitos de redundância de dados e reduz a probabilidade de uma operação de escalabilidade paralisar.
- O daemon de gerenciamento de logs foi atualizado para identificar todos os logs que estão em uso ativo com identificadores de arquivos abertos no armazenamento da instância local e nos processos associados. Essa atualização garante que o Amazon EMR exclua adequadamente os arquivos e recupere o espaço de armazenamento depois que os logs são arquivados no Amazon S3.

- Esta versão inclui um aprimoramento do daemon de gerenciamento de logs que exclui diretórios de etapas vazios e não utilizados no sistema de arquivos de cluster local. Um número excessivamente grande de diretórios vazios pode degradar a performance dos daemons do Amazon EMR e resultar na utilização excessiva do disco.
- Esta versão corrige um problema que pode ocorrer quando você cria um nó de borda ao replicar um dos nós primários de um cluster com vários nós primários. O nó de borda replicado pode causar atrasos nas operações de redução da escala verticalmente ou resultar em alta utilização de memória nos nós primários. Para obter mais informações sobre como criar um nó de borda para se comunicar com seu cluster EMR, consulte Edge Node Creator no `aws-samples` repositório em. GitHub
- Esta versão melhora o processo de automação que o Amazon EMR usa para remontar volumes do Amazon EBS em uma instância após uma reinicialização.
- Esta versão corrige um problema que resultou em lacunas intermitentes nas métricas do Hadoop que o Amazon EMR publica na Amazon. CloudWatch
- Esta versão corrige um problema com clusters do EMR em que uma atualização no arquivo de configuração do YARN que contém a lista de exclusão de nós do cluster é interrompida devido à utilização excessiva do disco. A atualização incompleta impede futuras operações de redução da escala verticalmente do cluster. Esta versão garante que o cluster permaneça íntegro e que as operações de escalabilidade funcionem conforme esperado.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Europa (Londres) , Ásia-Pacífico (Melbourn e), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (Oregon) AWS GovCloud , Oeste dos EUA (Norte da Califórnia), Ásia-Pací fico (AWS GovCloud Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv), Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha) , Europa (Estocolm o), Europa (Zurique) , Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Ásia-Pací fico (Jacarta) Ásia-Pacífico (Melbourne), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central), Oeste do Canadá (Calgary) AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud , (Leste dos EUA) AWS GovCloud , (Oeste dos EUA), China (Pequim), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Centro Leste (Emirados Árabes Unidos), Ásia-Pacífico (Melbourne),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Israel (Tel Aviv), Oeste do Canadá (Calgary), Europa (Zurique)
2.0.2024 001.0	4.14.352	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 816.0	4.14.350	21 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 809.0	4.14.349	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 719.0	4.14.348	25 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 223.0	4.14.336	8 de março de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 131.0	4.14.336	14 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 124.0	4.14.336	7 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad) Abad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 109.0	4.14.334	24 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 218.0	4.14.330	2 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 206.0	4.14.330	22 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 116.0	4.14.328	11 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 101.0	4.14.327	16 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 020.1	4.14.326	7 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 012.1	4.14.326	26 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 926.0	4.14.322	19 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 906.0	4.14.322	4 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 822.0	4.14.322	30 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 808.0	4.14.320	24 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 727.0	4.14.320	14 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central)

Versões de componente 6.10.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a

versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *CommunityVersion*-amzn-*EmrVersion*. O *EmrVersion* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.2	SDK do Amazon SageMaker Spark
delta	2.2.0	O Delta Lake é um formato de tabela aberto para grandes conjuntos de dados analíticos
emr-ddb	4.16.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.3.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.7.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.7.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.24.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	2.3.0	EMR S3Select Connector

Componente	Versão	Descrição
emr-wal-cli	1.0.0	Cli usada para lista/exclusão de emrwal.
emrfs	2.55.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.16.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.16.0	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	3.3.3-amzn-2.1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.

Componente	Versão	Descrição
hadoop-hdfs-datanode	3.3.3-amzn-2.1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.3.3-amzn-2.1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.3.3-amzn-2.1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	3.3.3-amzn-2.1	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.3.3-amzn-2.1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.3.3-amzn-2.1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	3.3.3-amzn-2.1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.3.3-amzn-2.1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	3.3.3-amzn-2.1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.

Componente	Versão	Descrição
hadoop-yarn-timeline-server	3.3.3-amzn-2.1	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	2.4.15-amzn-0.1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	2.4.15-amzn-0.1	Serviço para atender uma ou mais HBase regiões.
hbase-client	2.4.15-amzn-0.1	HBase cliente de linha de comando.
hbase-rest-server	2.4.15-amzn-0.1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.4.15-amzn-0.1	Serviço que fornece um endpoint Thrift para. HBase
hbase-operator-tools	2.4.15-amzn-0.1	Ferramenta de reparo para HBase clusters Apache.
hcatalog-client	3.1.3-amzn-3.1	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.3-amzn-3.1	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	3.1.3-amzn-3.1	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.3-amzn-3.1	O cliente da linha de comando do Hive.
hive-hbase	3.1.3-amzn-3.1	Cliente do Hive-hbase.
hive-metastore-server	3.1.3-amzn-3.1	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.3-amzn-3.1	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.12.2-amzn-0	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.12.2-amzn-0	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-trino	0.12.2-amzn-0	Biblioteca de pacotes para execução do Presto com o Hudi.
hudi-spark	0.12.2-amzn-0	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.10.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop

Componente	Versão	Descrição
iceberg	1.1.0-amzn-0	Apache Iceberg é um formato de tabela aberta para conjuntos de dados analíticos imensos
jupyterhub	1.5.0	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.1-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mxnet	1.9.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.
nvidia-cuda	11.8.0	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.5.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	5.1.2	As bibliotecas do phoenix para servidor e cliente

Componente	Versão	Descrição
phoenix-connectors	6.0.0-SNAPSHOT	Conectores do Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.278.1-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.278.1-amzn-0	O serviço que executa partes de uma consulta.
presto-client	0.278.1-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
trino-coordinator	403-amzn-0	Serviço que aceita consultas e gerencia a execução de consultas entre trino-workers.
trino-worker	403-amzn-0	O serviço que executa partes de uma consulta.

Componente	Versão	Descrição
trino-client	403-amzn-0	Cliente de linha de comando do Trino instalado em mestres de espera de um cluster de alta disponibilidade (HA) em que o servidor do Trino não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.0.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	3.3.1-amzn-0.1	Os clientes da linha de comando do Spark.
spark-history-server	3.3.1-amzn-0.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.3.1-amzn-0.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.3.1-amzn-0.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
spark-rapids	22.12.0-amzn-0	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs

Componente	Versão	Descrição
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.11.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.10.2-amzn-1.1	O aplicativo e as bibliotecas do YARN para Tez.
tez-on-worker	0.10.2-amzn-1.1	A aplicação e as bibliotecas do YARN nós de processam ento.
webserver	2.4.41+	O servidor Apache HTTP.
zeppelin-server	0.10.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.5.10	O serviço centralizado de manutenção de informações de configuração, nomenclat ura, fornecimento de sincroniz ação distribuída, e fornecime nto de serviços de grupo.
zookeeper-client	3.5.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 6.10.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-6.10.1

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.

Classificações	Descrição	Ações de reconfiguração
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.

Classificações	Descrição	Ações de reconfiguração
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.

Classificações	Descrição	Ações de reconfiguração
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classificações	Descrição	Ações de reconfiguração
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.

Classificações	Descrição	Ações de reconfiguração
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.	Not available.
iceberg-defaults	Alterar os valores no arquivo iceberg-defaults.conf do Iceberg.	Not available.
delta-defaults	Alterar os valores no arquivo delta-defaults.conf do Delta.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.

Classificações	Descrição	Ações de reconfiguração
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j2	Alterar as configurações de log4j2.properties do Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.

Classificações	Descrição	Ações de reconfiguração
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)

Classificações	Descrição	Ações de reconfiguração
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-lakeformation	Alterar os valores no arquivo lakeformation.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
trino-log	Alterar os valores no arquivo log.properties do Trino.	Restarts Trino-Server (for Trino)
trino-config	Alterar os valores no arquivo config.properties do Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Alterar os valores no arquivo password-authenticator.properties do Trino.	Restarts Trino-Server (for Trino)
trino-env	Alterar os valores no arquivo trino-env.sh do Trino.	Restarts Trino-Server (for Trino)
trino-node	Alterar os valores no arquivo node.properties do Trino.	Not available.

Classificações	Descrição	Ações de reconfiguração
trino-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Trino.	Not available.
trino-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Trino.	Not available.
trino-connector-delta	Alterar os valores no arquivo delta.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Alterar os valores no arquivo hive.properties do Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Alterar os valores no arquivo exchange-manager.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Alterar os valores no arquivo iceberg.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Alterar os valores no arquivo hudi.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Alterar os valores no arquivo jmx.properties do Trino.	Not available.
trino-connector-kafka	Alterar os valores no arquivo kafka.properties do Trino.	Not available.
trino-connector-localfile	Alterar os valores no arquivo localfile.properties do Trino.	Not available.
trino-connector-memory	Alterar os valores no arquivo memory.properties do Trino.	Not available.
trino-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Trino.	Not available.

Classificações	Descrição	Ações de reconfiguração
trino-connector-mysql	Alterar os valores no arquivo mysql.properties do Trino.	Not available.
trino-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Trino.	Not available.
trino-connector-raptor	Alterar os valores no arquivo raptor.properties do Trino.	Not available.
trino-connector-redis	Alterar os valores no arquivo redis.properties do Trino.	Not available.
trino-connector-redshift	Alterar os valores no arquivo redshift.properties do Trino.	Not available.
trino-connector-tpch	Alterar os valores no arquivo tpch.properties do Trino.	Not available.
trino-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Trino.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Alterar os valores no arquivo kms-logback.xml do Ranger KMS.	Not available.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
spark	As configurações administr adas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j2	Alterar os valores no arquivo log4j2.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie and HiveServe r2.

Classificações	Descrição	Ações de reconfiguração
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Log de alterações 6.10.1

Log de alterações da versão 6.10.1 e das notas da versão

Data	Event	Descrição
2023-08-30	Atualizar notas da versão	Adicionadas várias correções relacionadas ao ambiente de gerenciamento às notas da versão
2023-08-21	Publicação de documentos	Notas da versão 6.10.1 do Amazon EMR publicadas pela primeira vez
2023-08-16	Implantação concluída	Amazon EMR 6.10.1 totalment e implantado em todas as regiões compatíveis
2023-08-04	Lançamento inicial	A versão 6.10.1 do Amazon EMR foi implantada pela primeira vez em regiões comerciais limitadas

Versão 6.10.0 do Amazon EMR

Versões da aplicação 6.10.0

Essa versão inclui os seguintes aplicativos: Delta, Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, SqoopTensorFlow, Tez, TrinoZeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR

- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-6.10.0	emr-6.9.1	emr-6.9.0	emr-6.8.1
AWS SDK for Java	1.12.397	1.12.170	1.12.170	1.12.170
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.2.0	2.1.0	2.1.0	-
Flink	1.16.0	1.15.2	1.15.2	1.15.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.15	2.4.13	2.4.13	2.4.12
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.2.1
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.12.2-amzn-0	0.12.1-amzn-0	0.12.1-amzn-0	0.11.1-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.10.0
Iceberg	1.1.0-amzn-0	0.14.1-amzn-0	0.14.1-amzn-0	0.14.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.1.0

	emr-6.10.0	emr-6.9.1	emr-6.9.0	emr-6.8.1
JupyterHub	1.5.0	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,278	0,276	0,276	0,273
Spark	3.3.1	3.3.0	3.3.0	3.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.10.0	2.10.0	2.9.1
Tez	0.10.2	0.10.2	0.10.2	0.9.2
Trino (PrestoSQL)	403	398	398	388
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

Notas da versão 6.10.0

As notas da versão a seguir incluem informações para a versão 6.10.0 do Amazon EMR. As alterações são referentes à versão 6.9.0. Para obter informações sobre o cronograma da versão, consulte o log de alterações.

Novos recursos

- O Amazon EMR 6.10.0 é compatível com Apache Spark 3.3.1, Apache Spark RAPIDS 22.12.0, CUDA 11.8.0, Apache Hudi 0.12.2-amzn-0, Apache Iceberg 1.1.0-amzn-0, Trino 403 e PrestoDB 0.278.1.
- O Amazon EMR 6.10.0 inclui um conector Trino-Hudi nativo que fornece acesso de leitura aos dados nas tabelas Hudi. Você pode ativar o conector com `trino-cli --catalog hudi` e configurar o conector de acordo com suas necessidades com `trino-connector-hudi`. A integração nativa com o Amazon EMR significa que você não precisa mais usar `trino-connector-hive` para consultar tabelas do Hudi. Para obter uma lista das configurações compatíveis com o novo conector, consulte a página do conector do Hudi na documentação do Trino.
- O Amazon EMR 6.10.0 e versões posteriores oferecem suporte à integração do Apache Zeppelin com o Apache Flink. Consulte Usar trabalhos do Flink pelo Zeppelin no Amazon EMR para obter mais informações.

Problemas conhecidos

- O Hadoop 3.3.3 introduziu uma alteração no YARN (YARN-9608) que mantém os nós em que os contêineres eram executados em um estado de desativação até que a aplicação seja concluída. Essa alteração garante que dados locais, como dados embaralhados, não sejam perdidos e que você não precise executar o trabalho novamente. Essa abordagem também pode levar à subutilização de recursos em clusters com ou sem o ajuste de escala gerenciado habilitado.

  Para contornar esse problema no Amazon EMR 6.10.0, você pode definir o valor de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` como `false` em `yarn-site.xml`. Nas versões 6.11.0 e superiores do Amazon EMR, além das versões 6.8.1, 6.9.1 e 6.10.1, o config é definido como `false` por padrão para resolver esse problema.

Alterações, melhorias e problemas resolvidos

- O Amazon EMR 6.10.0 remove a dependência de `minimal-json.jar` para a integração do Amazon Redshift para Apache Spark e adiciona automaticamente os jars necessários relacionados ao Spark-Redshift ao caminho de classe do executor para o Spark: `spark-redshift.jar`, `spark-avro.jar` e `RedshiftJDBC.jar`.

- A versão 6.10.0 aprimora o daemon de gerenciamento de logs no cluster para monitorar pastas de log adicionais no cluster do EMR. Essa melhoria minimiza os cenários de utilização excessiva do disco.
- A versão 6.10.0 reinicia automaticamente o daemon de gerenciamento de logs no cluster quando ele é interrompido. Essa melhoria reduz o risco de os nós parecerem não íntegros devido à utilização excessiva do disco.
- O Amazon EMR 6.10.0 é compatível com endpoints regionais para mapeamento de usuários do EMRFS.
- O tamanho padrão do volume raiz aumentou para 15 GB nas versões 6.10.0 e superiores do Amazon EMR. O tamanho padrão do volume raiz das versões anteriores é de 10 GB.
- A versão 6.10.0 corrige um problema que fazia com que os trabalhos do Spark paralisassem quando todos os executores restantes do Spark estivessem em um host em desativação com o gerenciador de recursos do YARN.
- Com o Amazon EMR 6.6.0 a 6.9.x, as consultas INSERT com partição dinâmica e uma cláusula ORDER BY ou SORT BY sempre terá dois redutores. Esse problema é causado pela alteração do OSS HIVE-20703, que coloca a otimização da partição dinâmica de classificação sob uma decisão baseada em custos. Se sua workload não exigir a classificação de partições dinâmicas, recomendamos que você defina a propriedade `hive.optimize.sort.dynamic.partition.threshold` como `-1` para desabilitar o novo atributo e obter o número de redutores calculado corretamente. Esse problema foi corrigido no OSS Hive como parte do HIVE-22269 e foi corrigido no Amazon EMR 6.10.0.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

> **Note**
>
> Essa versão não recebe mais atualizações automáticas da AMI, pois foi substituída por uma ou mais versões de patch. A versão de patch é indicada pelo número após o segundo ponto decimal (`6.8.1`). Para ver se você está usando a versão de patch mais recente, verifique as versões disponíveis no Guia de lançamento ou verifique o menu suspenso de versões do Amazon EMR quando criar um cluster no console ou use a ação de API `ListReleaseLabels` ou da CLI `list-release-labels`. Para obter atualizações sobre novas versões, assine o feed RSS na página Novidades.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Europa (Londres) , Ásia-Pacífico (Melbourn e), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (Oregon) AWS GovCloud , Oeste dos EUA (Norte da Califórnia), Ásia-Pací fico (AWS GovCloud Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha) , Europa (Estocolm o), Europa (Zurique) , Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Ásia-Pací fico (Jacarta) Ásia-Pacífico (Melbourne), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central), Oeste do Canadá (Calgary) AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud , (Leste dos EUA) AWS GovCloud , (Oeste dos EUA), China (Pequim), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Centro Leste (Emirados Árabes Unidos), Ásia-Pacífico (Melbourne),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Israel (Tel Aviv), Oeste do Canadá (Calgary), Europa (Zurique)
2.0.2024 001.0	4.14.352	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 816.0	4.14.350	21 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 809.0	4.14.349	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 719.0	4.14.348	25 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)
2.0.2023 808.0	4.14.320	24 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), Israel (Tel Aviv).

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 727.0	4.14.320	14 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 719.0	4.14.320	2 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 628.0	4.14.318	12 de julho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Zurique), Europa (Milão), Europa (Espanha) , Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 612.0	4.14.314	23 de junho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Zurique), Europa (Milão), Europa (Espanha) , Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 504.1	4.14.313	16 de maio de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Zurique), Europa (Milão), Europa (Espanha) , Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 418.0	4.14.311	3 de maio de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Zurique), Europa (Milão), Europa (Espanha), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 404.1	4.14.311	18 de abril de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos)
2.0.2023 404.0	4.14.311	10 de abril de 2023	Leste dos EUA (Norte da Virgínia), Europa (Paris)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 320.0	4.14.309	30 de março de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 207.0	4.14.304	22 de fevereiro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos)

Versões de componente 6.10.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`.

Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *CommunityVersion*-amzn-*EmrVersion*. O *EmrVersion* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.2	SDK do Amazon SageMaker Spark
delta	2.2.0	O Delta Lake é um formato de tabela aberto para grandes conjuntos de dados analíticos
emr-ddb	4.16.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.3.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.7.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.7.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.24.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	2.3.0	EMR S3Select Connector

Componente	Versão	Descrição
emr-wal-cli	1.0.0	Cli usada para lista/exclusão de emrwal.
emrfs	2.55.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.16.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.16.0	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	3.3.3-amzn-2	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.

Componente	Versão	Descrição
hadoop-hdfs-datanode	3.3.3-amzn-2	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.3.3-amzn-2	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.3.3-amzn-2	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	3.3.3-amzn-2	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.3.3-amzn-2	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.3.3-amzn-2	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	3.3.3-amzn-2	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.3.3-amzn-2	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	3.3.3-amzn-2	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.

Componente	Versão	Descrição
hadoop-yarn-timeline-server	3.3.3-amzn-2	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	2.4.15-amzn-0	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	2.4.15-amzn-0	Serviço para atender uma ou mais HBase regiões.
hbase-client	2.4.15-amzn-0	HBase cliente de linha de comando.
hbase-rest-server	2.4.15-amzn-0	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.4.15-amzn-0	Serviço que fornece um endpoint Thrift para. HBase
hbase-operator-tools	2.4.15-amzn-0	Ferramenta de reparo para HBase clusters Apache.
hcatalog-client	3.1.3-amzn-3	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.3-amzn-3	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	3.1.3-amzn-3	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.3-amzn-3	O cliente da linha de comando do Hive.
hive-hbase	3.1.3-amzn-3	Cliente do Hive-hbase.
hive-metastore-server	3.1.3-amzn-3	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.3-amzn-3	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.12.2-amzn-0	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.12.2-amzn-0	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-trino	0.12.2-amzn-0	Biblioteca de pacotes para execução do Presto com o Hudi.
hudi-spark	0.12.2-amzn-0	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.10.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop

Componente	Versão	Descrição
iceberg	1.1.0-amzn-0	Apache Iceberg é um formato de tabela aberta para conjuntos de dados analíticos imensos
jupyterhub	1.5.0	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.1-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mxnet	1.9.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.
nvidia-cuda	11.8.0	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.5.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	5.1.2	As bibliotecas do phoenix para servidor e cliente

Componente	Versão	Descrição
phoenix-connectors	6.0.0-SNAPSHOT	Conectores do Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.278.1-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.278.1-amzn-0	O serviço que executa partes de uma consulta.
presto-client	0.278.1-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
trino-coordinator	403-amzn-0	Serviço que aceita consultas e gerencia a execução de consultas entre trino-workers.
trino-worker	403-amzn-0	O serviço que executa partes de uma consulta.

Componente	Versão	Descrição
trino-client	403-amzn-0	Cliente de linha de comando do Trino instalado em mestres de espera de um cluster de alta disponibilidade (HA) em que o servidor do Trino não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.0.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	3.3.1-amzn-0	Os clientes da linha de comando do Spark.
spark-history-server	3.3.1-amzn-0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.3.1-amzn-0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.3.1-amzn-0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
spark-rapids	22.12.0-amzn-0	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs

Componente	Versão	Descrição
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.11.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.10.2-amzn-1	O aplicativo e as bibliotecas do YARN para Tez.
tez-on-worker	0.10.2-amzn-1	A aplicação e as bibliotecas do YARN nós de processam ento.
webserver	2.4.41+	O servidor Apache HTTP.
zeppelin-server	0.10.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.5.10	O serviço centralizado de manutenção de informações de configuração, nomenclat ura, fornecimento de sincroniz ação distribuída, e fornecime nto de serviços de grupo.
zookeeper-client	3.5.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 6.10.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-6.10.0

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.

Classificações	Descrição	Ações de reconfiguração
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.

Classificações	Descrição	Ações de reconfiguração
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.

Classificações	Descrição	Ações de reconfiguração
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classificações	Descrição	Ações de reconfiguração
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.

Classificações	Descrição	Ações de reconfiguração
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.	Not available.
iceberg-defaults	Alterar os valores no arquivo iceberg-defaults.conf do Iceberg.	Not available.
delta-defaults	Alterar os valores no arquivo delta-defaults.conf do Delta.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.

Classificações	Descrição	Ações de reconfiguração
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j2	Alterar as configurações de log4j2.properties do Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.

Classificações	Descrição	Ações de reconfiguração
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)

Classificações	Descrição	Ações de reconfiguração
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-lakeformation	Alterar os valores no arquivo lakeformation.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
trino-log	Alterar os valores no arquivo log.properties do Trino.	Restarts Trino-Server (for Trino)
trino-config	Alterar os valores no arquivo config.properties do Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Alterar os valores no arquivo password-authenticator.properties do Trino.	Restarts Trino-Server (for Trino)
trino-env	Alterar os valores no arquivo trino-env.sh do Trino.	Restarts Trino-Server (for Trino)
trino-node	Alterar os valores no arquivo node.properties do Trino.	Not available.

Classificações	Descrição	Ações de reconfiguração
trino-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Trino.	Not available.
trino-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Trino.	Not available.
trino-connector-delta	Alterar os valores no arquivo delta.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Alterar os valores no arquivo hive.properties do Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Alterar os valores no arquivo exchange-manager.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Alterar os valores no arquivo iceberg.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Alterar os valores no arquivo hudi.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Alterar os valores no arquivo jmx.properties do Trino.	Not available.
trino-connector-kafka	Alterar os valores no arquivo kafka.properties do Trino.	Not available.
trino-connector-localfile	Alterar os valores no arquivo localfile.properties do Trino.	Not available.
trino-connector-memory	Alterar os valores no arquivo memory.properties do Trino.	Not available.
trino-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Trino.	Not available.

Classificações	Descrição	Ações de reconfiguração
trino-connector-mysql	Alterar os valores no arquivo mysql.properties do Trino.	Not available.
trino-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Trino.	Not available.
trino-connector-raptor	Alterar os valores no arquivo raptor.properties do Trino.	Not available.
trino-connector-redis	Alterar os valores no arquivo redis.properties do Trino.	Not available.
trino-connector-redshift	Alterar os valores no arquivo redshift.properties do Trino.	Not available.
trino-connector-tpch	Alterar os valores no arquivo tpch.properties do Trino.	Not available.
trino-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Trino.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Alterar os valores no arquivo kms-logback.xml do Ranger KMS.	Not available.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
spark	As configurações administr adas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j2	Alterar os valores no arquivo log4j2.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie and HiveServe r2.

Classificações	Descrição	Ações de reconfiguração
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Log de alterações 6.10.0

Log de alterações da versão 6.10.0 e das notas da versão

Data	Event	Descrição
2023-08-21	Atualizar	Adicionado um problema conhecido introduzido com o Hadoop 3.3.3.
2023-07-26	Atualizar	Novos rótulos de versão `2.0.20230612.0` e `2.0.20230628.0` do sistema operacional.
2023-03-02	Implantação concluída	Amazon EMR 6.10 totalment e implantado em todas as regiões compatíveis
2023-03-02	Publicação de documentos	Notas da versão 6.10 do Amazon EMR publicadas pela primeira vez
2023-02-27	Lançamento inicial	Versão 6.10 do Amazon EMR implantada em regiões comerciais limitadas

Versão 6.9.1 do Amazon EMR

Versões 6.9.1 da aplicação

Essa versão inclui os seguintes aplicativos: Delta, Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, SqoopTensorFlow, Tez, TrinoZeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-6.9.1	emr-6.9.0	emr-6.8.1	emr-6.8.0
AWS SDK for Java	1.12.170	1.12.170	1.12.170	1.12.170
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.1.0	2.1.0	-	-
Flink	1.15.2	1.15.2	1.15.1	1.15.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.13	2.4.13	2.4.12	2.4.12
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.2.1	3.2.1
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.12.1-amzn-0	0.12.1-amzn-0	0.11.1-amzn-0	0.11.1-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.10.0

	emr-6.9.1	emr-6.9.0	emr-6.8.1	emr-6.8.0
Iceberg	0.14.1-amzn-0	0.14.1-amzn-0	0.14.0-amzn-0	0.14.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,276	0,276	0,273	0,273
Spark	3.3.0	3.3.0	3.3.0	3.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.10.0	2.10.0	2.9.1	2.9.1
Tez	0.10.2	0.10.2	0.9.2	0.9.2
Trino (PrestoSQL)	398	398	388	388
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

Notas da versão 6.9.1

As notas da versão a seguir incluem informações para a versão 6.9.1 do Amazon EMR. As alterações são referentes à versão 6.9.0. Para obter informações sobre o cronograma da versão, consulte o Log de alterações 6.9.1.

Alterações, melhorias e problemas resolvidos

- O Hadoop 3.3.3 introduziu uma alteração no YARN (YARN-9608) que mantém os nós em que os contêineres eram executados em um estado de desativação até que a aplicação seja concluída. Essa alteração garante que dados locais, como dados embaralhados, não sejam perdidos e que você não precise executar o trabalho novamente. Essa abordagem também pode levar à subutilização de recursos em clusters com ou sem o ajuste de escala gerenciado habilitado.

  Com as versões 6.11.0 e superiores do Amazon EMR, além das versões 6.8.1, 6.9.1 e 6.10.1, o valor de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` é definido como `false` em `yarn-site.xml` para resolver esse problema.

  Embora a correção resolva os problemas introduzidos pelo YARN-9608, ela pode fazer com que os trabalhos do Hive falhem devido à perda de dados embaralhados em clusters com ajuste de escala gerenciado habilitado. Reduzimos esse risco nesta versão também ao configurar `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data` para workloads do Hive. Essa configuração só está disponível com as versões 6.11.0 e superiores do Amazon EMR.
- O coletor de métricas não enviará métricas para o ambiente de gerenciamento após o failover do nó primário em clusters com a configuração de grupos de instâncias.
- Essa versão elimina novas tentativas de solicitações HTTP com falha nos endpoints do coletor de métricas.
- Esta versão inclui uma alteração que permite que clusters de alta disponibilidade se recuperem de um estado de falha após a reinicialização.
- Esta versão corrige um problema em que grandes exceções de estouro criadas UIDs por usuários causavam exceções de estouro.
- Esta versão corrige problemas de tempo limite com o processo de reconfiguração do Amazon EMR.
- Esta versão inclui correções de segurança.
- Esta versão corrige um problema em que clusters que estão executando workloads no Spark com o Amazon EMR podem receber silenciosamente resultados incorretos com `contains`,

`startsWith`, `endsWith` e `like`. Esse problema ocorre quando você usa as expressões em campos particionados que têm metadados no Hive3 Metastore Server (HMS) no Amazon EMR.

- Com o Amazon EMR 6.6.0 a 6.9.x, as consultas INSERT com partição dinâmica e uma cláusula ORDER BY ou SORT BY sempre terá dois redutores. Esse problema é causado pela alteração do OSS HIVE-20703, que coloca a otimização da partição dinâmica de classificação sob uma decisão baseada em custos. Se sua workload não exigir a classificação de partições dinâmicas, recomendamos que você defina a propriedade `hive.optimize.sort.dynamic.partition.threshold` como `-1` para desabilitar o novo atributo e obter o número de redutores calculado corretamente. Esse problema foi corrigido no OSS Hive como parte do HIVE-22269 e foi corrigido no Amazon EMR 6.10.0.
- O Hive pode sofrer perda de dados quando você usa o HDFS como um diretório temporário e habilitou a mesclagem de arquivos pequenos e a tabela contém caminhos de partição estáticos.
- Esta versão corrige um problema de performance com o Hive se a mesclagem de arquivos pequenos (desabilitada por padrão) é habilitada no final da tarefa de ETL.
- Esta versão corrige um problema com controle de utilização no lado do Glue quando não há funções definidas pelo usuário (UDF).
- Esta versão corrige um problema que exclui logs de contêineres pelo serviço de agregação de logs de nó antes que o pusher de logs possa enviá-los para o S3 em caso de desativação do YARN.
- Esta versão corrige o manuseio de arquivos compactados/arquivados com rastreamento persistente de arquivos de armazenamento para. HBase
- Esta versão corrige um problema que afetou a performance do Spark quando você definiu um valor de `true` padrão para a configuração `spark.yarn.heterogeneousExecutors.enabled` no `spark-defaults.conf`.
- Esta versão corrige um problema com a falha do Reduce Task em ler dados embaralhados. O problema causou falhas na consulta do Hive com um erro de memória corrompida.
- Esta versão corrige um problema que fazia com que o provisionador de nós falhasse se o serviço HDFS NameNode (NN) ficasse preso no modo de segurança durante a substituição do nó.
- Esta versão adiciona um mecanismo de nova tentativa ao fluxo de trabalho de escalabilidade de clusters para clusters do EMR que executam o Presto ou o Trino. Essa melhoria reduz o risco de que o redimensionamento do cluster fique paralisado indefinidamente devido a uma única falha na operação de redimensionamento. Ela também aprimora a utilização dos clusters, porque seu cluster aumenta e reduz a escala verticalmente com mais rapidez.
- Esta versão melhora a lógica de redução da escala verticalmente do cluster para que o cluster não tente reduzir a escala verticalmente dos nós centrais abaixo da configuração do fator de replicação

do HDFS para o cluster. Isso se alinha aos seus requisitos de redundância de dados e reduz a probabilidade de uma operação de escalabilidade paralisar.

- O daemon de gerenciamento de logs foi atualizado para identificar todos os logs que estão em uso ativo com identificadores de arquivos abertos no armazenamento da instância local e nos processos associados. Essa atualização garante que o Amazon EMR exclua adequadamente os arquivos e recupere o espaço de armazenamento depois que os logs são arquivados no Amazon S3.
- Esta versão inclui um aprimoramento do daemon de gerenciamento de logs que exclui diretórios de etapas vazios e não utilizados no sistema de arquivos de cluster local. Um número excessivamente grande de diretórios vazios pode degradar a performance dos daemons do Amazon EMR e resultar na utilização excessiva do disco.
- Esta versão corrige um problema que pode ocorrer quando você cria um nó de borda ao replicar um dos nós primários de um cluster com vários nós primários. O nó de borda replicado pode causar atrasos nas operações de redução da escala verticalmente ou resultar em alta utilização de memória nos nós primários. Para obter mais informações sobre como criar um nó de borda para se comunicar com seu cluster EMR, consulte Edge Node Creator no `aws-samples` repositório em. GitHub
- Esta versão melhora o processo de automação que o Amazon EMR usa para remontar volumes do Amazon EBS em uma instância após uma reinicialização.
- Esta versão corrige um problema que resultou em lacunas intermitentes nas métricas do Hadoop que o Amazon EMR publica na Amazon. CloudWatch
- Esta versão corrige um problema com clusters do EMR em que uma atualização no arquivo de configuração do YARN que contém a lista de exclusão de nós do cluster é interrompida devido à utilização excessiva do disco. A atualização incompleta impede futuras operações de redução da escala verticalmente do cluster. Esta versão garante que o cluster permaneça íntegro e que as operações de escalabilidade funcionem conforme esperado.
- Esta versão aprimora o daemon de gerenciamento de logs no cluster para monitorar pastas de log adicionais no cluster do EMR. Essa melhoria minimiza os cenários de utilização excessiva do disco.
- Esta versão reinicia automaticamente o daemon de gerenciamento de logs no cluster quando ele é interrompido. Essa melhoria reduz o risco de os nós parecerem não íntegros devido à utilização excessiva do disco.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do

Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Europa (Londres) , Ásia-Pacífico (Melbourn e), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (Oregon) AWS GovCloud , Oeste dos EUA (Norte da Califórnia), Ásia-Pací fico (AWS GovCloud Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha) , Europa (Estocolm o), Europa (Zurique) , Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Ásia-Pací fico (Jacarta) Ásia-Pacífico (Melbourne), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central), Oeste do Canadá (Calgary) AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud , (Leste dos EUA) AWS GovCloud , (Oeste dos EUA), China (Pequim), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Centro Leste (Emirados Árabes Unidos), Ásia-Pacífico (Melbourne),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Israel (Tel Aviv), Oeste do Canadá (Calgary), Europa (Zurique)
2.0.2024 001.0	4.14.352	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 816.0	4.14.350	21 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 809.0	4.14.349	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 719.0	4.14.348	25 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 223.0	4.14.336	8 de março de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central) , Israel (Tel Aviv), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), AWS GovCloud China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 131.0	4.14.336	14 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central) , Israel (Tel Aviv), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), AWS GovCloud China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 124.0	4.14.336	7 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central) , Israel (Tel Aviv), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), AWS GovCloud China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 109.0	4.14.334	24 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central) , Israel (Tel Aviv), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), AWS GovCloud China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 218.0	4.14.330	2 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central) , Israel (Tel Aviv), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), AWS GovCloud China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 206.0	4.14.330	22 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central) , Israel (Tel Aviv), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), AWS GovCloud China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 116.0	4.14.328	11 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central) , Israel (Tel Aviv), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), AWS GovCloud China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 101.0	4.14.327	16 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central) , Israel (Tel Aviv), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), AWS GovCloud China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 020.1	4.14.326	7 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central) , Israel (Tel Aviv), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), AWS GovCloud China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 012.1	4.14.326	26 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central) , Israel (Tel Aviv), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), AWS GovCloud China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 926.0	4.14.322	19 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central) , Israel (Tel Aviv), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), AWS GovCloud China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 906.0	4.14.322	4 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pací fico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central), Israel (Tel Aviv).

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 822.0	4.14.322	30 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pací fico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central), Israel (Tel Aviv).

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 808.0	4.14.320	24 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pací fico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central), Israel (Tel Aviv).

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 727.0	4.14.320	14 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central)

Versões de componente 6.9.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a

versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *CommunityVersion*-amzn-*EmrVersion*. O *EmrVersion* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.2	SDK do Amazon SageMaker Spark
delta	2.1.0	O Delta Lake é um formato de tabela aberto para grandes conjuntos de dados analíticos
emr-ddb	4.16.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.3.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.6.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.7.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.23.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	2.2.0	EMR S3Select Connector

Componente	Versão	Descrição
emrfs	2.54.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.15.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.15.2	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	3.3.3-amzn-1.1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	3.3.3-amzn-1.1	O serviço de nível de nó do HDFS para armazenamento de blocos.

Componente	Versão	Descrição
hadoop-hdfs-library	3.3.3-amzn-1.1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.3.3-amzn-1.1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	3.3.3-amzn-1.1	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.3.3-amzn-1.1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.3.3-amzn-1.1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	3.3.3-amzn-1.1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.3.3-amzn-1.1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	3.3.3-amzn-1.1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hadoop-yarn-timeline-server	3.3.3-amzn-1.1	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.

Componente	Versão	Descrição
hbase-hmaster	2.4.13-amzn-0.1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	2.4.13-amzn-0.1	Serviço para atender uma ou mais HBase regiões.
hbase-client	2.4.13-amzn-0.1	HBase cliente de linha de comando.
hbase-rest-server	2.4.13-amzn-0.1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.4.13-amzn-0.1	Serviço que fornece um endpoint Thrift para. HBase
hbase-operator-tools	2.4.13-amzn-0.1	Ferramenta de reparo para HBase clusters Apache.
hcatalog-client	3.1.3-amzn-2.1	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.3-amzn-2.1	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	3.1.3-amzn-2.1	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.

Componente	Versão	Descrição
hive-client	3.1.3-amzn-2.1	O cliente da linha de comando do Hive.
hive-hbase	3.1.3-amzn-2.1	Cliente do Hive-hbase.
hive-metastore-server	3.1.3-amzn-2.1	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.3-amzn-2.1	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.12.1-amzn-0	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.12.1-amzn-0	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-trino	0.12.1-amzn-0	Biblioteca de pacotes para execução do Presto com o Hudi.
hudi-spark	0.12.1-amzn-0	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.10.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop

Componente	Versão	Descrição
iceberg	0.14.1-amzn-0	Apache Iceberg é um formato de tabela aberta para conjuntos de dados analíticos imensos
jupyterhub	1.4.1	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.1-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mxnet	1.9.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.
nvidia-cuda	11.7.0	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.5.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	5.1.2	As bibliotecas do phoenix para servidor e cliente

Componente	Versão	Descrição
phoenix-connectors	6.0.0-SNAPSHOT	Conectores do Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.276-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.276-amzn-0	O serviço que executa partes de uma consulta.
presto-client	0.276-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
trino-coordinator	398-amzn-0	Serviço que aceita consultas e gerencia a execução de consultas entre trino-workers.
trino-worker	398-amzn-0	O serviço que executa partes de uma consulta.

Componente	Versão	Descrição
trino-client	398-amzn-0	Cliente de linha de comando do Trino instalado em mestres de espera de um cluster de alta disponibilidade (HA) em que o servidor do Trino não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.0.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	3.3.0-amzn-1.1	Os clientes da linha de comando do Spark.
spark-history-server	3.3.0-amzn-1.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.3.0-amzn-1.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.3.0-amzn-1.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
spark-rapids	22.08.0-amzn-0	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs

Componente	Versão	Descrição
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.10.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.10.2-amzn-0.1	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.41+	O servidor Apache HTTP.
zeppelin-server	0.10.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.5.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.5.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 6.9.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as

classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-6.9.1

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode,

Classificações	Descrição	Ações de reconfiguração
		SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.

Classificações	Descrição	Ações de reconfiguração
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.

Classificações	Descrição	Ações de reconfiguração
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classificações	Descrição	Ações de reconfiguração
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.

Classificações	Descrição	Ações de reconfiguração
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.	Not available.
iceberg-defaults	Alterar os valores no arquivo iceberg-defaults.conf do Iceberg.	Not available.
delta-defaults	Alterar os valores no arquivo delta-defaults.conf do Delta.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.

Classificações	Descrição	Ações de reconfiguração
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j2	Alterar as configurações de log4j2.properties do Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.

Classificações	Descrição	Ações de reconfiguração
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)

Classificações	Descrição	Ações de reconfiguração
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-lakeformation	Alterar os valores no arquivo lakeformation.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
trino-log	Alterar os valores no arquivo log.properties do Trino.	Restarts Trino-Server (for Trino)
trino-config	Alterar os valores no arquivo config.properties do Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Alterar os valores no arquivo password-authenticator.properties do Trino.	Restarts Trino-Server (for Trino)
trino-env	Alterar os valores no arquivo trino-env.sh do Trino.	Restarts Trino-Server (for Trino)
trino-node	Alterar os valores no arquivo node.properties do Trino.	Not available.

Classificações	Descrição	Ações de reconfiguração
trino-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Trino.	Not available.
trino-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Trino.	Not available.
trino-connector-delta	Alterar os valores no arquivo delta.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Alterar os valores no arquivo hive.properties do Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Alterar os valores no arquivo exchange-manager.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Alterar os valores no arquivo iceberg.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Alterar os valores no arquivo jmx.properties do Trino.	Not available.
trino-connector-kafka	Alterar os valores no arquivo kafka.properties do Trino.	Not available.
trino-connector-localfile	Alterar os valores no arquivo localfile.properties do Trino.	Not available.
trino-connector-memory	Alterar os valores no arquivo memory.properties do Trino.	Not available.
trino-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Trino.	Not available.
trino-connector-mysql	Alterar os valores no arquivo mysql.properties do Trino.	Not available.

Classificações	Descrição	Ações de reconfiguração
trino-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Trino.	Not available.
trino-connector-raptor	Alterar os valores no arquivo raptor.properties do Trino.	Not available.
trino-connector-redis	Alterar os valores no arquivo redis.properties do Trino.	Not available.
trino-connector-redshift	Alterar os valores no arquivo redshift.properties do Trino.	Not available.
trino-connector-tpch	Alterar os valores no arquivo tpch.properties do Trino.	Not available.
trino-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Trino.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.

Classificações	Descrição	Ações de reconfiguração
spark	As configurações administr adas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j2	Alterar os valores no arquivo log4j2.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie and HiveServe r2.

Classificações	Descrição	Ações de reconfiguração
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Log de alterações 6.9.1

Log de alterações da versão 6.9.1 e das notas da versão

Data	Event	Descrição
2023-08-30	Atualizar notas da versão	Adicionadas várias correções relacionadas ao ambiente de gerenciamento às notas da versão
2023-08-21	Publicação de documentos	Notas da versão 6.9.1 do Amazon EMR publicadas pela primeira vez
2023-08-16	Implantação concluída	Amazon EMR 6.9.1 totalmente implantado em todas as regiões compatíveis
2023-08-04	Lançamento inicial	A versão 6.9.1 do Amazon EMR foi implantada pela primeira vez em regiões comerciais limitadas

Versão 6.9.0 do Amazon EMR

Versões da aplicação 6.9.0

Essa versão inclui os seguintes aplicativos: Delta, Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, SqoopTensorFlow, Tez, TrinoZeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR

- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-6.9.0	emr-6.8.1	emr-6.8.0	emr-6.7.0
AWS SDK for Java	1.12.170	1.12.170	1.12.170	1.12.170
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.1.0	-	-	-
Flink	1.15.2	1.15.1	1.15.1	1.14.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.13	2.4.12	2.4.12	2.4.4
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.2.1	3.2.1	3.2.1
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.12.1-amzn-0	0.11.1-amzn-0	0.11.1-amzn-0	0.11.0-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.10.0
Iceberg	0.14.1-amzn-0	0.14.0-amzn-0	0.14.0-amzn-0	0.13.1-amzn-0
JupyterEnterpriseGateway	2.6.0	2.1.0	2.1.0	2.1.0

	emr-6.9.0	emr-6.8.1	emr-6.8.0	emr-6.7.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.8.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,276	0,273	0,273	0,272
Spark	3.3.0	3.3.0	3.3.0	3.2.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.10.0	2.9.1	2.9.1	2.4.1
Tez	0.10.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	398	388	388	378
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.0
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.7

Notas da versão 6.9.0

As notas da versão a seguir incluem informações para a versão 6.9.0 do Amazon EMR. As alterações são referentes à versão 6.8.0 do Amazon EMR. Para obter informações sobre o cronograma da versão, consulte o log de alterações.

Novos atributos

- A versão 6.9.0 do Amazon EMR é compatível com Apache Spark RAPIDS 22.08.0, Apache Hudi 0.12.1, Apache Iceberg 0.14.1, Trino 398 e Tez 0.10.2.
- A versão 6.9.0 do Amazon EMR inclui uma nova aplicação de código aberto, Delta Lake 2.1.0.
- A integração do Amazon Redshift para Apache Spark está inclusa nas versões 6.9.0 e posteriores do Amazon EMR. Anteriormente uma ferramenta de código aberto, a integração nativa é um conector do Spark que você pode usar para criar aplicações do Apache Spark que realizam a leitura e a gravação de dados no Amazon Redshift e no Amazon Redshift sem servidor. Para obter mais informações, consulte Usar a integração do Amazon Redshift para Apache Spark com o Amazon EMR .
- A versão 6.9.0 do Amazon EMR adiciona suporte ao arquivamento de logs no Amazon S3 durante a redução da escala do cluster verticalmente. Anteriormente, só era possível arquivar arquivos de log no Amazon S3 durante o encerramento do cluster. A nova capacidade garante que os arquivos de log gerados no cluster persistam no Amazon S3 mesmo após o encerramento do nó. Para obter mais informações, consulte Configurar registro em log e depuração do cluster.
- Para dar suporte a consultas de longa execução, o Trino agora inclui um mecanismo de execução tolerante a falhas. A execução tolerante a falhas atenua as falhas nas consultas ao tentar novamente as consultas com falha ou as tarefas dos seus componentes.
- Você pode usar o Apache Flink no Amazon EMR para `BATCH` unificado e processamento de `STREAM` de tabelas do Apache Hive ou metadados de qualquer tablesource do Flink, como Iceberg, Kinesis ou Kafka. Você pode especificar o AWS Glue Data Catalog como metastore para o Flink usando a API AWS Management Console,, AWS CLI ou Amazon EMR. Para obter mais informações, consulte Configurar o Flink no Amazon EMR.
- Agora você pode especificar funções de tempo de execução AWS Identity and Access Management (IAM) e controle de acesso AWS Lake Formation baseado para consultas do Apache Spark, Apache Hive e Presto no Amazon EC2 EMR em clusters com o Amazon AI Studio. SageMaker Para obter mais informações, consulte Configurar funções de runtime para as etapas do Amazon EMR.

Problemas conhecidos

- Para a versão 6.9.0 do Amazon EMR, o Trino não funciona em clusters habilitados para o Apache Ranger. Se você precisar usar o Trino com o Ranger, entre em contato com o Suporte.

- Se você usar a integração do Amazon Redshift para Apache Spark e tiver um time, timetz, timestamp ou timestamptz com precisão de microssegundos no formato Parquet, o conector arredondará os valores de tempo para o valor de milissegundo mais próximo. Como solução alternativa, use o parâmetro `unload_s3_format` do formato de descarregamento de texto.
- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
  - O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

  Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.
- As conexões com clusters do Amazon EMR do Amazon SageMaker AI Studio podem falhar intermitentemente com um código de resposta 403 Forbidden. Esse erro ocorre quando a configuração do perfil do IAM no cluster leva mais de 60 segundos. Como solução alternativa, você pode instalar um patch do Amazon EMR para permitir novas tentativas e aumentar o tempo limite para um mínimo de 300 segundos. Use as etapas a seguir para aplicar a ação de bootstrap quando iniciar o cluster.

  1. Baixe o script de bootstrap e os arquivos RPM do Amazon URIs S3 a seguir.

     ```
     s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/gcsc/replace-rpms.sh
     s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/gcsc/emr-secret-agent-1.18.0-SNAPSHOT20221121212949.noarch.rpm
     ```

  2. Carregue os arquivos da etapa anterior em um bucket do Amazon S3 de sua propriedade. O bucket deve estar no mesmo Região da AWS local em que você planeja iniciar o cluster.

3. Inclua a seguinte ação de bootstrap ao iniciar o cluster do EMR. Substitua *bootstrap_URI* e *RPM_URI* pelo correspondente URIs do Amazon S3.

```
--bootstrap-actions "Path=bootstrap_URI,Args=[RPM_URI]"
```

- Com as versões 5.36.0 e 6.6.0 a 6.9.0 do Amazon EMR, os componentes do serviço `SecretAgent` e `RecordServer` podem sofrer perda de dados de log devido a uma configuração incorreta do padrão de nome de arquivo nas propriedades do Log4j2. A configuração incorreta faz com que os componentes gerem somente um arquivo de log por dia. Quando a estratégia de rotação ocorre, ela substitui o arquivo existente em vez de gerar um novo arquivo de log, conforme esperado. Como solução alternativa, use uma ação de bootstrap para gerar arquivos de log a cada hora e acrescentar um número inteiro de incremento automático no nome do arquivo para lidar com a rotação.

  Para as versões 6.6.0 a 6.9.0 do Amazon EMR, use a seguinte ação de bootstrap ao iniciar um cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

  Para o Amazon EMR 5.36.0, use a ação de bootstrap a seguir ao iniciar um cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

- O Apache Flink fornece FileSystem conectores nativos S3 FileSystem e Hadoop, que permitem que os aplicativos criem FileSink e gravem os dados no Amazon S3. Isso FileSink falha com uma das duas exceções a seguir.

```
java.lang.UnsupportedOperationException: Recoverable writers on Hadoop are only
 supported for HDFS
```

```
Caused by: java.lang.NoSuchMethodError:
 org.apache.hadoop.io.retry.RetryPolicies.retryOtherThanRemoteAndSaslException(Lorg/
apache/hadoop/io/retry/RetryPolicy;Ljava/util/Map;)Lorg/apache/hadoop/io/retry/
RetryPolicy;
```

```
                                        at
 org.apache.hadoop.yarn.client.RMProxy.createRetryPolicy(RMProxy.java:302) ~[hadoop-
yarn-common-3.3.3-amzn-0.jar:?]
```

Como solução alternativa, você pode instalar um patch do Amazon EMR, que corrige o problema acima no Flink. Para aplicar a ação de bootstrap quando iniciar o cluster, execute as etapas a seguir.

1. Baixe o flink-rpm no bucket Amazon S3. Seu caminho de RPM é `s3://DOC-EXAMPLE-BUCKET/rpms/flink/`.
2. Baixe o script de bootstrap e os arquivos RPM do Amazon S3 usando o URI a seguir. *regionName*Substitua pelo Região da AWS local em que você planeja iniciar o cluster.

   ```
   s3://emr-data-access-control-regionName/customer-bootstrap-actions/gcsc/replace-
   rpms.sh
   ```
3. O Hadoop 3.3.3 introduziu uma alteração no YARN (YARN-9608) que mantém os nós em que os contêineres eram executados em um estado de desativação até que a aplicação seja concluída. Essa alteração garante que dados locais, como dados embaralhados, não sejam perdidos e que você não precise executar o trabalho novamente. Nas versões 6.8.0 e 6.9.0 do Amazon EMR, essa abordagem também pode levar à subutilização de recursos em clusters com ou sem o ajuste de escala gerenciado habilitado.

   Com o Amazon EMR 6.10.0, há uma solução alternativa para esse problema: definir o valor de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` como `false` em `yarn-site.xml`. Nas versões 6.11.0 e superiores do Amazon EMR, além das versões 6.8.1, 6.9.1 e 6.10.1, o config é definido como `false` por padrão para resolver esse problema.

Alterações, melhorias e problemas resolvidos

- Para as versões 6.9.0 e posteriores do Amazon EMR, todos os componentes instalados pelo Amazon EMR que usam bibliotecas do Log4j usam o Log4j versão 2.17.1 ou posterior.
- Ao usar o conector DynamoDB com o Spark nas versões 6.6.0, 6.7.0 e 6.8.0 do Amazon EMR, todas as leituras da tabela retornam um resultado vazio, mesmo que a divisão de entrada faça referência a dados que não estão vazios. A versão 6.9.0 do Amazon EMR corrige esse problema.
- O Amazon EMR 6.9.0 adiciona suporte limitado ao controle de acesso baseado no Lake Formation com o Apache Hudi ao ler dados usando o Spark SQL. O suporte se destina a consultas SELECT

usando o Spark SQL e é limitado ao controle de acesso em nível de coluna. Para obter mais informações, consulte Hudi e Lake Formation.

- Quando você usa o Amazon EMR 6.9.0 para criar um cluster do Hadoop com Rótulos de nós habilitados, a API de métricas do YARN retorna informações agregadas em todas as partições, em vez de na partição padrão. Para obter mais informações, consulte YARN-11414.
- Com a versão 6.9.0 do Amazon EMR, atualizamos o Trino para a versão 398, que usa Java 17. A versão anterior do Trino compatível com o Amazon EMR 6.8.0 era Trino 388 em execução no Java 11. Para obter mais informações sobre essa alteração, consulte Atualizações do Trino para Java 17 no blog do Trino.
- Esta versão corrige um problema de incompatibilidade de sequência de tempo entre o Apache BigTop e o Amazon EMR na EC2 sequência de inicialização do cluster. Essa incompatibilidade de sequência de tempo ocorre quando um sistema tenta realizar duas ou mais operações ao mesmo tempo em vez de fazê-las na sequência correta. Como resultado, determinadas configurações de cluster apresentaram tempos limite de inicialização da instância e tempos de inicialização do cluster mais lentos.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

> **Note**
>
> Essa versão não recebe mais atualizações automáticas da AMI, pois foi substituída por uma ou mais versões de patch. A versão de patch é indicada pelo número após o segundo ponto decimal (`6.8.1`). Para ver se você está usando a versão de patch mais recente, verifique as versões disponíveis no Guia de lançamento ou verifique o menu suspenso de versões do Amazon EMR quando criar um cluster no console ou use a ação de API `ListReleaseLabels` ou da CLI `list-release-labels`. Para obter atualizações sobre novas versões, assine o feed RSS na página Novidades.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Europa (Londres) , Ásia-Pacífico (Melbourn e), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (Oregon) AWS GovCloud , Oeste dos EUA (Norte da Califórnia), Ásia-Pací fico (AWS GovCloud Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha) , Europa (Estocolm o), Europa (Zurique) , Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Ásia-Pací fico (Jacarta) Ásia-Pacífico (Melbourne), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central), Oeste do Canadá (Calgary) AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud , (Leste dos EUA) AWS GovCloud , (Oeste dos EUA), China (Pequim), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Centro Leste (EAU), Ásia-Pacífico (Melbourne),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Israel (Tel Aviv), Europa (Zurique)
2.0.2024 001.0	4.14.352	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 816.0	4.14.350	21 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 809.0	4.14.349	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 719.0	4.14.348	25 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)
2.0.2023 808.0	4.14.320	24 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), Israel (Tel Aviv).

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 727.0	4.14.320	14 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 719.0	4.14.320	2 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 628.0	4.14.318	12 de julho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 612.0	4.14.314	23 de junho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 504.1	4.14.313	16 de maio de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 418.0	4.14.311	3 de maio de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 404.1	4.14.311	18 de abril de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)
2.0.2023 404.0	4.14.311	10 de abril de 2023	Leste dos EUA (Norte da Virgínia), Europa (Paris)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 320.0	4.14.309	30 de março de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 307.0	4.14.305	15 de março de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 207.0	4.14.304	22 de fevereiro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 210.1	4.14.301	12 de janeiro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 103.3	4.14.296	5 de dezembro de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

Versões de componente 6.9.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a

versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.2	SDK do Amazon SageMaker Spark
delta	2.1.0	O Delta Lake é um formato de tabela aberto para grandes conjuntos de dados analíticos
emr-ddb	4.16.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.3.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.6.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.7.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.23.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	2.2.0	EMR S3Select Connector

Componente	Versão	Descrição
emrfs	2.54.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.15.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.15.2	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	3.3.3-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	3.3.3-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.

Componente	Versão	Descrição
hadoop-hdfs-library	3.3.3-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.3.3-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	3.3.3-amzn-1	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.3.3-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.3.3-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	3.3.3-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.3.3-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	3.3.3-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hadoop-yarn-timeline-server	3.3.3-amzn-1	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.

Componente	Versão	Descrição
hbase-hmaster	2.4.13-amzn-0	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	2.4.13-amzn-0	Serviço para atender uma ou mais HBase regiões.
hbase-client	2.4.13-amzn-0	HBase cliente de linha de comando.
hbase-rest-server	2.4.13-amzn-0	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.4.13-amzn-0	Serviço que fornece um endpoint Thrift para. HBase
hbase-operator-tools	2.4.13-amzn-0	Ferramenta de reparo para HBase clusters Apache.
hcatalog-client	3.1.3-amzn-2	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.3-amzn-2	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	3.1.3-amzn-2	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.

Componente	Versão	Descrição
hive-client	3.1.3-amzn-2	O cliente da linha de comando do Hive.
hive-hbase	3.1.3-amzn-2	Cliente do Hive-hbase.
hive-metastore-server	3.1.3-amzn-2	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.3-amzn-2	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.12.1-amzn-0	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.12.1-amzn-0	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-trino	0.12.1-amzn-0	Biblioteca de pacotes para execução do Presto com o Hudi.
hudi-spark	0.12.1-amzn-0	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.10.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop

Componente	Versão	Descrição
iceberg	0.14.1-amzn-0	Apache Iceberg é um formato de tabela aberta para conjuntos de dados analíticos imensos
jupyterhub	1.4.1	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.1-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mxnet	1.9.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.
nvidia-cuda	11.7.0	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.5.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	5.1.2	As bibliotecas do phoenix para servidor e cliente

Componente	Versão	Descrição
phoenix-connectors	6.0.0-SNAPSHOT	Conectores do Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.276-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.276-amzn-0	O serviço que executa partes de uma consulta.
presto-client	0.276-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
trino-coordinator	398-amzn-0	Serviço que aceita consultas e gerencia a execução de consultas entre trino-workers.
trino-worker	398-amzn-0	O serviço que executa partes de uma consulta.

Componente	Versão	Descrição
trino-client	398-amzn-0	Cliente de linha de comando do Trino instalado em mestres de espera de um cluster de alta disponibilidade (HA) em que o servidor do Trino não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.0.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	3.3.0-amzn-1	Os clientes da linha de comando do Spark.
spark-history-server	3.3.0-amzn-1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.3.0-amzn-1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.3.0-amzn-1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
spark-rapids	22.08.0-amzn-0	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs

Componente	Versão	Descrição
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.10.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.10.2-amzn-0	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.41+	O servidor Apache HTTP.
zeppelin-server	0.10.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.5.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.5.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 6.9.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as

classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-6.9.0

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode,

Classificações	Descrição	Ações de reconfiguração
		SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.

Classificações	Descrição	Ações de reconfiguração
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.

Classificações	Descrição	Ações de reconfiguração
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classificações	Descrição	Ações de reconfiguração
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.

Classificações	Descrição	Ações de reconfiguração
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.	Not available.
iceberg-defaults	Alterar os valores no arquivo iceberg-defaults.conf do Iceberg.	Not available.
delta-defaults	Alterar os valores no arquivo delta-defaults.conf do Delta.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.

Classificações	Descrição	Ações de reconfiguração
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j2	Alterar as configurações de log4j2.properties do Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.

Classificações	Descrição	Ações de reconfiguração
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)

Classificações	Descrição	Ações de reconfiguração
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-lakeformation	Alterar os valores no arquivo lakeformation.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
trino-log	Alterar os valores no arquivo log.properties do Trino.	Restarts Trino-Server (for Trino)
trino-config	Alterar os valores no arquivo config.properties do Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Alterar os valores no arquivo password-authenticator.properties do Trino.	Restarts Trino-Server (for Trino)
trino-env	Alterar os valores no arquivo trino-env.sh do Trino.	Restarts Trino-Server (for Trino)
trino-node	Alterar os valores no arquivo node.properties do Trino.	Not available.

Classificações	Descrição	Ações de reconfiguração
trino-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Trino.	Not available.
trino-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Trino.	Not available.
trino-connector-delta	Alterar os valores no arquivo delta.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Alterar os valores no arquivo hive.properties do Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Alterar os valores no arquivo exchange-manager.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Alterar os valores no arquivo iceberg.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Alterar os valores no arquivo jmx.properties do Trino.	Not available.
trino-connector-kafka	Alterar os valores no arquivo kafka.properties do Trino.	Not available.
trino-connector-localfile	Alterar os valores no arquivo localfile.properties do Trino.	Not available.
trino-connector-memory	Alterar os valores no arquivo memory.properties do Trino.	Not available.
trino-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Trino.	Not available.
trino-connector-mysql	Alterar os valores no arquivo mysql.properties do Trino.	Not available.

Classificações	Descrição	Ações de reconfiguração
trino-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Trino.	Not available.
trino-connector-raptor	Alterar os valores no arquivo raptor.properties do Trino.	Not available.
trino-connector-redis	Alterar os valores no arquivo redis.properties do Trino.	Not available.
trino-connector-redshift	Alterar os valores no arquivo redshift.properties do Trino.	Not available.
trino-connector-tpch	Alterar os valores no arquivo tpch.properties do Trino.	Not available.
trino-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Trino.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.

Classificações	Descrição	Ações de reconfiguração
spark	As configurações administr adas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j2	Alterar os valores no arquivo log4j2.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie and HiveServe r2.

Classificações	Descrição	Ações de reconfiguração
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Log de alterações 6.9.0

Log de alterações da versão 6.9.0 e das notas da versão

Data	Event	Descrição
2023-08-30	Atualizar notas da versão	Adicionada correção para um problema de incompatibilidade de sequência de tempo
2023-08-21	Atualizar notas da versão	Adicionado um problema conhecido com o Hadoop 3.3.3.
2023-07-26	Atualizar	Novos rótulos de versão `2.0.20230612.0` e `2.0.20230628.0` do sistema operacional.
2022-12-13	Notas da versão atualizadas	Recurso adicionado e problema conhecido para tempo de execução com SageMaker IA
2022-11-29	Notas da versão e documentação atualizadas	Adicionado atributo para a integração do Amazon Redshift para Apache Spark
2022-11-23	Notas da versão atualizadas	Removida entrada Log4j
2022-11-18	Implantação concluída	Amazon EMR 6.9 totalmente implantado em todas as regiões compatíveis
2022-11-18	Publicação de documentos	Notas da versão 6.9 do Amazon EMR publicadas pela primeira vez

Data	Event	Descrição
2022-11-14	Lançamento inicial	Versão 6.9 do Amazon EMR implantada em regiões comerciais limitadas

Versão 6.8.1 do Amazon EMR

Versões 6.8.1 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, SqoopTensorFlow, Tez, Trino, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-6.8.1	emr-6.8.0	emr-6.7.0	emr-6.6.0
AWS SDK for Java	1.12.170	1.12.170	1.12.170	1.12.170
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.10
AmazonCloudWatchAgent	-	-	-	-

	emr-6.8.1	emr-6.8.0	emr-6.7.0	emr-6.6.0
Delta	-	-	-	-
Flink	1.15.1	1.15.1	1.14.2	1.14.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.12	2.4.12	2.4.4	2.4.4
HCatalog	3.1.3	3.1.3	3.1.3	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.3	3.1.3	3.1.3	3.1.2
Hudi	0.11.1-amzn-0	0.11.1-amzn-0	0.11.0-amzn-0	0.10.1-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.10.0
Iceberg	0.14.0-amzn-0	0.14.0-amzn-0	0.13.1-amzn-0	0.13.1
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.8.0	1.8.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,273	0,273	0,272	0,267
Spark	3.3.0	3.3.0	3.2.1	3.2.0

	emr-6.8.1	emr-6.8.0	emr-6.7.0	emr-6.6.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.9.1	2.9.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	388	388	378	367
Zeppelin	0.10.1	0.10.1	0.10.0	0.10.0
ZooKeeper	3.5.10	3.5.10	3.5.7	3.5.7

Notas da versão 6.8.1

As notas da versão a seguir incluem informações para a versão 6.8.1 do Amazon EMR. As alterações são referentes à versão 6.8.0. Para obter informações sobre o cronograma da versão, consulte o Log de alterações 6.8.1.

Alterações, melhorias e problemas resolvidos

- O Hadoop 3.3.3 introduziu uma alteração no YARN (YARN-9608) que mantém os nós em que os contêineres eram executados em um estado de desativação até que a aplicação seja concluída. Essa alteração garante que dados locais, como dados embaralhados, não sejam perdidos e que você não precise executar o trabalho novamente. Essa abordagem também pode levar à subutilização de recursos em clusters com ou sem o ajuste de escala gerenciado habilitado.

  Com as versões 6.11.0 e superiores do Amazon EMR, além das versões 6.8.1, 6.9.1 e 6.10.1, o valor de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` é definido como `false` em `yarn-site.xml` para resolver esse problema.

  Embora a correção resolva os problemas introduzidos pelo YARN-9608, ela pode fazer com que os trabalhos do Hive falhem devido à perda de dados embaralhados em clusters com ajuste de escala gerenciado habilitado. Reduzimos esse risco nesta versão também ao configurar `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data` para workloads do Hive. Essa configuração só está disponível com as versões 6.11.0 e superiores do Amazon EMR.

- O coletor de métricas não enviará métricas para o ambiente de gerenciamento após o failover do nó primário em clusters com a configuração de grupos de instâncias.
- Essa versão elimina novas tentativas de solicitações HTTP com falha nos endpoints do coletor de métricas.
- Esta versão inclui uma alteração que permite que clusters de alta disponibilidade se recuperem de um estado de falha após a reinicialização.
- Esta versão corrige um problema em que grandes exceções de estouro criadas UIDs por usuários causavam exceções de estouro.
- Esta versão corrige problemas de tempo limite com o processo de reconfiguração do Amazon EMR.
- Esta versão evita um problema em que a falha na reconfiguração pode interromper outros processos não relacionados.
- Esta versão inclui correções de segurança.
- Esta versão corrige um problema em que clusters que estão executando workloads no Spark com o Amazon EMR podem receber silenciosamente resultados incorretos com `contains`, `startsWith`, `endsWith` e `like`. Esse problema ocorre quando você usa as expressões em campos particionados que têm metadados no Hive3 Metastore Server (HMS) no Amazon EMR.
- Com o Amazon EMR 6.6.0 a 6.9.x, as consultas INSERT com partição dinâmica e uma cláusula ORDER BY ou SORT BY sempre terá dois redutores. Esse problema é causado pela alteração do OSS HIVE-20703, que coloca a otimização da partição dinâmica de classificação sob uma decisão baseada em custos. Se sua workload não exigir a classificação de partições dinâmicas, recomendamos que você defina a propriedade `hive.optimize.sort.dynamic.partition.threshold` como `-1` para desabilitar o novo atributo e obter o número de redutores calculado corretamente. Esse problema foi corrigido no OSS Hive como parte do HIVE-22269 e foi corrigido no Amazon EMR 6.10.0.
- O Hive pode sofrer perda de dados quando você usa o HDFS como um diretório temporário e habilitou a mesclagem de arquivos pequenos e a tabela contém caminhos de partição estáticos.
- Esta versão corrige um problema de performance com o Hive se a mesclagem de arquivos pequenos (desabilitada por padrão) é habilitada no final da tarefa de ETL.
- Esta versão corrige um problema com controle de utilização no lado do Glue quando não há funções definidas pelo usuário (UDF).
- Esta versão corrige um problema que exclui logs de contêineres pelo serviço de agregação de logs de nó antes que o pusher de logs possa enviá-los para o S3 em caso de desativação do YARN.

- Esta versão corrige o manuseio de arquivos compactados/arquivados com rastreamento persistente de arquivos de armazenamento para. HBase
- Esta versão corrige um problema que afetou a performance do Spark quando você definiu um valor de `true` padrão para a configuração `spark.yarn.heterogeneousExecutors.enabled` no `spark-defaults.conf`.
- Esta versão corrige um problema com a falha do Reduce Task em ler dados embaralhados. O problema causou falhas na consulta do Hive com um erro de memória corrompida.
- Esta versão corrige um problema que fazia com que o provisionador de nós falhasse se o serviço HDFS NameNode (NN) ficasse preso no modo de segurança durante a substituição do nó.
- Esta versão adiciona um mecanismo de nova tentativa ao fluxo de trabalho de escalabilidade de clusters para clusters do EMR que executam o Presto ou o Trino. Essa melhoria reduz o risco de que o redimensionamento do cluster fique paralisado indefinidamente devido a uma única falha na operação de redimensionamento. Ela também aprimora a utilização dos clusters, porque seu cluster aumenta e reduz a escala verticalmente com mais rapidez.
- Esta versão melhora a lógica de redução da escala verticalmente do cluster para que o cluster não tente reduzir a escala verticalmente dos nós centrais abaixo da configuração do fator de replicação do HDFS para o cluster. Isso se alinha aos seus requisitos de redundância de dados e reduz a probabilidade de uma operação de escalabilidade paralisar.
- O daemon de gerenciamento de logs foi atualizado para identificar todos os logs que estão em uso ativo com identificadores de arquivos abertos no armazenamento da instância local e nos processos associados. Essa atualização garante que o Amazon EMR exclua adequadamente os arquivos e recupere o espaço de armazenamento depois que os logs são arquivados no Amazon S3.
- Esta versão inclui um aprimoramento do daemon de gerenciamento de logs que exclui diretórios de etapas vazios e não utilizados no sistema de arquivos de cluster local. Um número excessivamente grande de diretórios vazios pode degradar a performance dos daemons do Amazon EMR e resultar na utilização excessiva do disco.
- Esta versão corrige um problema que pode ocorrer quando você cria um nó de borda ao replicar um dos nós primários de um cluster com vários nós primários. O nó de borda replicado pode causar atrasos nas operações de redução da escala verticalmente ou resultar em alta utilização de memória nos nós primários. Para obter mais informações sobre como criar um nó de borda para se comunicar com seu cluster EMR, consulte Edge Node Creator no `aws-samples` repositório em. GitHub
- Esta versão melhora o processo de automação que o Amazon EMR usa para remontar volumes do Amazon EBS em uma instância após uma reinicialização.

- Esta versão corrige um problema que resultou em lacunas intermitentes nas métricas do Hadoop que o Amazon EMR publica na Amazon. CloudWatch
- Esta versão corrige um problema com clusters do EMR em que uma atualização no arquivo de configuração do YARN que contém a lista de exclusão de nós do cluster é interrompida devido à utilização excessiva do disco. A atualização incompleta impede futuras operações de redução da escala verticalmente do cluster. Esta versão garante que o cluster permaneça íntegro e que as operações de escalabilidade funcionem conforme esperado.
- Esta versão aprimora o daemon de gerenciamento de logs no cluster para monitorar pastas de log adicionais no cluster do EMR. Essa melhoria minimiza os cenários de utilização excessiva do disco.
- Esta versão reinicia automaticamente o daemon de gerenciamento de logs no cluster quando ele é interrompido. Essa melhoria reduz o risco de os nós parecerem não íntegros devido à utilização excessiva do disco.
- Esta versão adiciona suporte ao arquivamento de logs no Amazon S3 durante a redução da escala do cluster verticalmente. Anteriormente, só era possível arquivar arquivos de log no Amazon S3 durante o encerramento do cluster. A nova capacidade garante que os arquivos de log gerados no cluster persistam no Amazon S3 mesmo após o encerramento do nó. Para obter mais informações, consulte Configurar registro em log e depuração do cluster.
- Esta versão corrige um problema que ocorreu quando o URI do Amazon S3 para uma ação de bootstrap terminava com um número de porta, por exemplo: `a.b.c.d:4345`. O Amazon EMR os estava analisando incorretamente URIs, então qualquer ação de bootstrap associada falharia.
- Esta versão corrige um problema de incompatibilidade de sequência de tempo entre o Apache BigTop e o Amazon EMR na EC2 sequência de inicialização do cluster. Essa incompatibilidade de sequência de tempo ocorre quando um sistema tenta realizar duas ou mais operações ao mesmo tempo em vez de fazê-las na sequência correta. Como resultado, determinadas configurações de cluster apresentaram tempos limite de inicialização da instância e tempos de inicialização do cluster mais lentos.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Europa (Londres) , Ásia-Pacífico (Melbourn e), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (Oregon) AWS GovCloud , Oeste dos EUA (Norte da Califórnia), Ásia-Pací fico (AWS GovCloud Cingapura), Ásia-Pací fico (Sydney), Israel (Tel

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Aviv), Canadá (Central) , Europa (Espanha), China (Ningxia), Europa (Zurique)
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Hyderabad), Ásia-Pacífico (Jacarta) Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , AWS GovCloud (Leste dos EUA), AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Oeste dos EUA), China (Pequim), China (Ningxia), Israel (Tel Aviv)
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), China (Pequim), China (Ningxia), Oriente Médio (Emirados Árabes Unidos), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 001.0	4.14.352	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 816.0	4.14.350	21 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 809.0	4.14.349	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 719.0	4.14.348	25 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 223.0	4.14.336	8 de março de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central) , Israel (Tel Aviv), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), AWS GovCloud China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 131.0	4.14.336	14 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central) , Israel (Tel Aviv), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), AWS GovCloud China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 124.0	4.14.336	7 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central) , Israel (Tel Aviv), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), AWS GovCloud China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 109.0	4.14.334	24 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central) , Israel (Tel Aviv), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), AWS GovCloud China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 218.0	4.14.330	2 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central) , Israel (Tel Aviv), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), AWS GovCloud China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 206.0	4.14.330	22 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central) , Israel (Tel Aviv), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), AWS GovCloud China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 116.0	4.14.328	11 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central) , Israel (Tel Aviv), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), AWS GovCloud China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 101.0	4.14.327	16 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central) , Israel (Tel Aviv), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), AWS GovCloud China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 020.1	4.14.326	7 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central) , Israel (Tel Aviv), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), AWS GovCloud China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 012.1	4.14.326	26 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central) , Israel (Tel Aviv), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), AWS GovCloud China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 926.0	4.14.322	19 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central) , Israel (Tel Aviv), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), AWS GovCloud China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 906.0	4.14.322	4 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pací fico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 822.0	4.14.322	30 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pací fico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 808.0	4.14.320	24 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pací fico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 727.0	4.14.320	14 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central)

Versões de componente 6.8.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a

versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *CommunityVersion*-amzn-*EmrVersion*. O *EmrVersion* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.2	SDK do Amazon SageMaker Spark
emr-ddb	4.16.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.2.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.7.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.22.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	2.1.0	EMR S3Select Connector
emrfs	2.53.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.

Componente	Versão	Descrição
flink-client	1.15.1	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.15.1	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	3.2.1-amzn-8.1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	3.2.1-amzn-8.1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.2.1-amzn-8.1	O cliente de linha de comando e biblioteca do HDFS

Componente	Versão	Descrição
hadoop-hdfs-namenode	3.2.1-amzn-8.1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	3.2.1-amzn-8.1	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.2.1-amzn-8.1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.2.1-amzn-8.1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-8.1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.2.1-amzn-8.1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	3.2.1-amzn-8.1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	3.2.1-amzn-8.1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.

Componente	Versão	Descrição
hbase-hmaster	2.4.12-amzn-0.1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	2.4.12-amzn-0.1	Serviço para atender uma ou mais HBase regiões.
hbase-client	2.4.12-amzn-0.1	HBase cliente de linha de comando.
hbase-rest-server	2.4.12-amzn-0.1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.4.12-amzn-0.1	Serviço que fornece um endpoint Thrift para. HBase
hbase-operator-tools	2.4.12-amzn-0.1	Ferramenta de reparo para HBase clusters Apache.
hcatalog-client	3.1.3-amzn-1.1	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.3-amzn-1.1	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	3.1.3-amzn-1.1	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.

Componente	Versão	Descrição
hive-client	3.1.3-amzn-1.1	O cliente da linha de comando do Hive.
hive-hbase	3.1.3-amzn-1.1	Cliente do Hive-hbase.
hive-metastore-server	3.1.3-amzn-1.1	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.3-amzn-1.1	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.11.1-amzn-0	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.11.1-amzn-0	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-trino	0.11.1-amzn-0	Biblioteca de pacotes para execução do Presto com o Hudi.
hudi-spark	0.11.1-amzn-0	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.10.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop

Componente	Versão	Descrição
iceberg	0.14.0-amzn-0	Apache Iceberg é um formato de tabela aberta para conjuntos de dados analíticos imensos
jupyterhub	1.4.1	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.1-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mxnet	1.9.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.
nvidia-cuda	11.7.0	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.5.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	5.1.2	As bibliotecas do phoenix para servidor e cliente

Componente	Versão	Descrição
phoenix-connectors	5.1.2	Conectores do Apache Phoenix para Spark-3
phoenix-query-server	5.1.2	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.273.3-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.273.3-amzn-0	O serviço que executa partes de uma consulta.
presto-client	0.273.3-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
trino-coordinator	388-amzn-0	Serviço que aceita consultas e gerencia a execução de consultas entre trino-workers.
trino-worker	388-amzn-0	O serviço que executa partes de uma consulta.

Componente	Versão	Descrição
trino-client	388-amzn-0	Cliente de linha de comando do Trino instalado em mestres de espera de um cluster de alta disponibilidade (HA) em que o servidor do Trino não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.0.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	3.3.0-amzn-0.1	Os clientes da linha de comando do Spark.
spark-history-server	3.3.0-amzn-0.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.3.0-amzn-0.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.3.0-amzn-0.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
spark-rapids	22.06.0-amzn-0	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs

Componente	Versão	Descrição
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.9.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.41+	O servidor Apache HTTP.
zeppelin-server	0.10.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.5.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.5.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 6.8.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as

classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-6.8.1

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode,

Classificações	Descrição	Ações de reconfiguração
		SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.

Classificações	Descrição	Ações de reconfiguração
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.

Classificações	Descrição	Ações de reconfiguração
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classificações	Descrição	Ações de reconfiguração
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.

Classificações	Descrição	Ações de reconfiguração
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.	Not available.
iceberg-defaults	Alterar os valores no arquivo iceberg-defaults.conf do Iceberg.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.

Classificações	Descrição	Ações de reconfiguração
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j	Alterar as configurações de log4j.properties no Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.

Classificações	Descrição	Ações de reconfiguração
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
trino-log	Alterar os valores no arquivo log.properties do Trino.	Restarts Trino-Server (for Trino)
trino-config	Alterar os valores no arquivo config.properties do Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Alterar os valores no arquivo password-authenticator.properties do Trino.	Restarts Trino-Server (for Trino)
trino-env	Alterar os valores no arquivo trino-env.sh do Trino.	Restarts Trino-Server (for Trino)
trino-node	Alterar os valores no arquivo node.properties do Trino.	Not available.
trino-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Trino.	Not available.
trino-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Trino.	Not available.
trino-connector-hive	Alterar os valores no arquivo hive.properties do Trino.	Restarts Trino-Server (for Trino)

Classificações	Descrição	Ações de reconfiguração
trino-connector-iceberg	Alterar os valores no arquivo iceberg.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Alterar os valores no arquivo jmx.properties do Trino.	Not available.
trino-connector-kafka	Alterar os valores no arquivo kafka.properties do Trino.	Not available.
trino-connector-localfile	Alterar os valores no arquivo localfile.properties do Trino.	Not available.
trino-connector-memory	Alterar os valores no arquivo memory.properties do Trino.	Not available.
trino-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Trino.	Not available.
trino-connector-mysql	Alterar os valores no arquivo mysql.properties do Trino.	Not available.
trino-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Trino.	Not available.
trino-connector-raptor	Alterar os valores no arquivo raptor.properties do Trino.	Not available.
trino-connector-redis	Alterar os valores no arquivo redis.properties do Trino.	Not available.
trino-connector-redshift	Alterar os valores no arquivo redshift.properties do Trino.	Not available.
trino-connector-tpch	Alterar os valores no arquivo tpch.properties do Trino.	Not available.
trino-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Trino.	Not available.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j2	Alterar os valores no arquivo log4j2.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.

Classificações	Descrição	Ações de reconfiguração
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie and HiveServe r2.
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.

Classificações	Descrição	Ações de reconfiguração
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Log de alterações 6.8.1

Log de alterações da versão 6.8.1 e das notas da versão

Data	Event	Descrição
2023-08-30	Atualizar notas da versão	Adicionadas várias correções relacionadas ao ambiente de gerenciamento às notas da versão
2023-08-21	Publicação de documentos	Notas da versão 6.8.1 do Amazon EMR publicadas pela primeira vez
2023-08-16	Implantação concluída	Amazon EMR 6.8.1 totalmente implantado em todas as regiões compatíveis
2023-08-04	Lançamento inicial	A versão 6.8.1 do Amazon EMR foi implantada pela primeira vez em regiões comerciais limitadas

Versão 6.8.0 do Amazon EMR

Versões 6.8.0 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, SqoopTensorFlow, Tez, Trino, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-6.8.0	emr-6.7.0	emr-6.6.0	emr-6.5.0
AWS SDK for Java	1.12.170	1.12.170	1.12.170	1.12.31
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.15.1	1.14.2	1.14.2	1.14.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.12	2.4.4	2.4.4	2.4.4
HCatalog	3.1.3	3.1.3	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.3	3.1.3	3.1.2	3.1.2
Hudi	0.11.1-amzn-0	0.11.0-amzn-0	0.10.1-amzn-0	0.9.0-amzn-1

	emr-6.8.0	emr-6.7.0	emr-6.6.0	emr-6.5.0
Hue	4.10.0	4.10.0	4.10.0	4.9.0
Iceberg	0.14.0-amzn-0	0.13.1-amzn-0	0.13.1	0.12.0
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.8.0	1.8.0	1.8.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,273	0,272	0,267	0,261
Spark	3.3.0	3.2.1	3.2.0	3.1.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.9.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	388	378	367	360
Zeppelin	0.10.1	0.10.0	0.10.0	0.10.0
ZooKeeper	3.5.10	3.5.7	3.5.7	3.5.7

Notas da versão 6.8.0

As notas da versão a seguir incluem informações para a versão 6.8.0 do Amazon EMR. As alterações são referentes à versão 6.7.0.

Novos atributos

- O recurso de etapas do Amazon EMR agora oferece suporte a endpoints e clientes Apache Livy. JDBC/ODBC Para obter mais informações, consulte Configurar funções de runtime para as etapas do Amazon EMR.
- A versão 6.8.0 do Amazon EMR vem com a versão 2.4.12 do Apache. HBase Com essa HBase versão, você pode arquivar e excluir suas HBase tabelas. O processamento de arquivos do Amazon S3 renomeia todos os arquivos da tabela para o diretório de arquivos. Isso pode ser um processo custoso e demorado. Agora, você pode pular o processamento de arquivos e rapidamente eliminar e excluir tabelas grandes. Para obter mais informações, consulte Usando a HBase concha.

Problemas conhecidos

- O Hadoop 3.3.3 introduziu uma alteração no YARN (YARN-9608) que mantém os nós em que os contêineres eram executados em um estado de desativação até que a aplicação seja concluída. Essa alteração garante que dados locais, como dados embaralhados, não sejam perdidos e que você não precise executar o trabalho novamente. Nas versões 6.8.0 e 6.9.0 do Amazon EMR, essa abordagem também pode levar à subutilização de recursos em clusters com ou sem o ajuste de escala gerenciado habilitado.

  Com o Amazon EMR 6.10.0, há uma solução alternativa para esse problema: definir o valor de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` como `false` em `yarn-site.xml`. Nas versões 6.11.0 e superiores do Amazon EMR, além das versões 6.8.1, 6.9.1 e 6.10.1, o config é definido como `false` por padrão para resolver esse problema.

Alterações, melhorias e problemas resolvidos

- Quando a versão 6.5.0, 6.6.0 ou 6.7.0 do Amazon EMR leu as tabelas do Apache Phoenix por meio do shell do Apache Spark, o Amazon EMR produziu um `NoSuchMethodError`. A versão 6.8.0 do Amazon EMR corrige esse problema.

- A versão 6.8.0 do Amazon EMR vem com o Apache Hudi 0.11.1; no entanto, os clusters do Amazon EMR 6.8.0 também são compatíveis com o código aberto `hudi-spark3.3-bundle_2.12` do Hudi 0.12.0.
- A versão 6.8.0 do Amazon EMR vem com a versão 3.3.0 do Apache Spark. Esta versão do Spark usa o Apache Log4j 2 e o arquivo `log4j2.properties` para configurar o Log4j nos processos do Spark. Se você usar o Spark no cluster ou criar clusters do EMR com parâmetros de configuração personalizados e quiser atualizar para a versão 6.8.0 do Amazon EMR, deverá migrar para a nova classificação de configuração `spark-log4j2` e para o formato de chave do Apache Log4j 2. Para obter mais informações, consulte Migrar do Apache Log4j 1.x para Log4j 2.x.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

> **Note**
>
> Essa versão não recebe mais atualizações automáticas da AMI, pois foi substituída por uma ou mais versões de patch. A versão de patch é indicada pelo número após o segundo ponto decimal (`6.8.`*`1`*). Para ver se você está usando a versão de patch mais recente, verifique as versões disponíveis no Guia de lançamento ou verifique o menu suspenso de versões do Amazon EMR quando criar um cluster no console ou use a ação de API `ListReleaseLabels` ou da CLI `list-release-labels`. Para obter atualizações sobre novas versões, assine o feed RSS na página Novidades.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Europa (Londres) , Ásia-Pacífico (Melbourn e), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (Oregon) AWS GovCloud , Oeste dos EUA (Norte da Califórnia), Ásia-Pací fico (AWS GovCloud Cingapura), Ásia-Pací fico (Sydney), Israel (Tel Aviv), Canadá (Central) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Hyderabad), Ásia-Pacífico (Jacarta) Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , AWS GovCloud (Leste dos EUA), AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Oeste dos EUA), China (Pequim), China (Ningxia), Israel (Tel Aviv)
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Centro Leste (EAU), Ásia-Pacífico (Melbourne),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Israel (Tel Aviv), Europa (Zurique)
2.0.2024 001.0	4.14.352	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 816.0	4.14.350	21 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 809.0	4.14.349	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 719.0	4.14.348	25 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv)
2.0.2023 808.0	4.14.320	24 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórni a), Oeste dos EUA (Oregon), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 727.0	4.14.320	14 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 719.0	4.14.320	2 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 628.0	4.14.318	12 de julho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 612.0	4.14.314	23 de junho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 504.1	4.14.313	16 de maio de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 418.0	4.14.311	3 de maio de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 404.1	4.14.311	18 de abril de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)
2.0.2023 404.0	4.14.311	10 de abril de 2023	Leste dos EUA (Norte da Virgínia), Europa (Paris)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 320.0	4.14.309	30 de março de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 307.0	4.14.305	15 de março de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 207.0	4.14.304	22 de fevereiro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 119.1	4.14.301	3 de fevereiro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 210.1	4.14.301	22 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 103.3	4.14.296	5 de dezembro de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 004.0	4.14.294	2 de novembro de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 912.1	4.14.291	6 de setembro de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

Problemas conhecidos

- Ao usar o conector DynamoDB com o Spark nas versões 6.6.0, 6.7.0 e 6.8.0 do Amazon EMR, todas as leituras da tabela retornam um resultado vazio, mesmo que a divisão de entrada faça referência a dados que não estão vazios. Isso ocorre porque o Spark 3.2.0 define `spark.hadoopRDD.ignoreEmptySplits` como `true` por padrão. Como solução alternativa,

defina explicitamente `spark.hadoopRDD.ignoreEmptySplits` como `false`. A versão 6.9.0 do Amazon EMR corrige esse problema.

- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
  - O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

  Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.
- Com as versões 5.36.0 e 6.6.0 a 6.9.0 do Amazon EMR, os componentes do serviço `SecretAgent` e `RecordServer` podem sofrer perda de dados de log devido a uma configuração incorreta do padrão de nome de arquivo nas propriedades do Log4j2. A configuração incorreta faz com que os componentes gerem somente um arquivo de log por dia. Quando a estratégia de rotação ocorre, ela substitui o arquivo existente em vez de gerar um novo arquivo de log, conforme esperado. Como solução alternativa, use uma ação de bootstrap para gerar arquivos de log a cada hora e acrescentar um número inteiro de incremento automático no nome do arquivo para lidar com a rotação.

  Para as versões 6.6.0 a 6.9.0 do Amazon EMR, use a seguinte ação de bootstrap ao iniciar um cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

  Para o Amazon EMR 5.36.0, use a ação de bootstrap a seguir ao iniciar um cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

Para obter mais informações sobre o cronograma da versão, consulte o log de alterações.

Versões de componente 6.8.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.2	SDK do Amazon SageMaker Spark
emr-ddb	4.16.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.2.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.7.0	Ambiente Conda para bloco de anotações do emr, que

Componente	Versão	Descrição
		inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.22.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	2.1.0	EMR S3Select Connector
emrfs	2.53.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.15.1	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.15.1	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.

Componente	Versão	Descrição
hadoop-client	3.2.1-amzn-8	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	3.2.1-amzn-8	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.2.1-amzn-8	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.2.1-amzn-8	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	3.2.1-amzn-8	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.2.1-amzn-8	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.2.1-amzn-8	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	3.2.1-amzn-8	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.2.1-amzn-8	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	3.2.1-amzn-8	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	3.2.1-amzn-8	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	2.4.12-amzn-0	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	2.4.12-amzn-0	Serviço para atender uma ou mais HBase regiões.
hbase-client	2.4.12-amzn-0	HBase cliente de linha de comando.
hbase-rest-server	2.4.12-amzn-0	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.4.12-amzn-0	Serviço que fornece um endpoint Thrift para. HBase
hbase-operator-tools	2.4.12-amzn-0	Ferramenta de reparo para HBase clusters Apache.
hcatalog-client	3.1.3-amzn-1	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.

Componente	Versão	Descrição
hcatalog-server	3.1.3-amzn-1	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	3.1.3-amzn-1	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.3-amzn-1	O cliente da linha de comando do Hive.
hive-hbase	3.1.3-amzn-1	Cliente do Hive-hbase.
hive-metastore-server	3.1.3-amzn-1	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.3-amzn-1	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.11.1-amzn-0	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.11.1-amzn-0	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-trino	0.11.1-amzn-0	Biblioteca de pacotes para execução do Presto com o Hudi.

Componente	Versão	Descrição
hudi-spark	0.11.1-amzn-0	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.10.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
iceberg	0.14.0-amzn-0	Apache Iceberg é um formato de tabela aberta para conjuntos de dados analíticos imensos
jupyterhub	1.4.1	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.1-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mxnet	1.9.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.
nvidia-cuda	11.7.0	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.

Componente	Versão	Descrição
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.5.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	5.1.2	As bibliotecas do phoenix para servidor e cliente
phoenix-connectors	5.1.2	Conectores do Apache Phoenix para Spark-3
phoenix-query-server	5.1.2	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.273.3-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.273.3-amzn-0	O serviço que executa partes de uma consulta.
presto-client	0.273.3-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
trino-coordinator	388-amzn-0	Serviço que aceita consultas e gerencia a execução de consultas entre trino-workers.

Componente	Versão	Descrição
trino-worker	388-amzn-0	O serviço que executa partes de uma consulta.
trino-client	388-amzn-0	Cliente de linha de comando do Trino instalado em mestres de espera de um cluster de alta disponibilidade (HA) em que o servidor do Trino não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.0.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	3.3.0-amzn-0	Os clientes da linha de comando do Spark.
spark-history-server	3.3.0-amzn-0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.3.0-amzn-0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.3.0-amzn-0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.

Componente	Versão	Descrição
spark-rapids	22.06.0-amzn-0	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.9.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.41+	O servidor Apache HTTP.
zeppelin-server	0.10.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.5.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.5.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 6.8.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-6.8.0

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.

Classificações	Descrição	Ações de reconfiguração
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.

Classificações	Descrição	Ações de reconfiguração
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.

Classificações	Descrição	Ações de reconfiguração
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classificações	Descrição	Ações de reconfiguração
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.

Classificações	Descrição	Ações de reconfiguração
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.	Not available.
iceberg-defaults	Alterar os valores no arquivo iceberg-defaults.conf do Iceberg.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.

Classificações	Descrição	Ações de reconfiguração
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j	Alterar as configurações de log4j.properties no Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.

Classificações	Descrição	Ações de reconfiguração
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
trino-log	Alterar os valores no arquivo log.properties do Trino.	Restarts Trino-Server (for Trino)
trino-config	Alterar os valores no arquivo config.properties do Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Alterar os valores no arquivo password-authenticator.properties do Trino.	Restarts Trino-Server (for Trino)
trino-env	Alterar os valores no arquivo trino-env.sh do Trino.	Restarts Trino-Server (for Trino)
trino-node	Alterar os valores no arquivo node.properties do Trino.	Not available.
trino-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Trino.	Not available.
trino-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Trino.	Not available.
trino-connector-hive	Alterar os valores no arquivo hive.properties do Trino.	Restarts Trino-Server (for Trino)

Classificações	Descrição	Ações de reconfiguração
trino-connector-iceberg	Alterar os valores no arquivo iceberg.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Alterar os valores no arquivo jmx.properties do Trino.	Not available.
trino-connector-kafka	Alterar os valores no arquivo kafka.properties do Trino.	Not available.
trino-connector-localfile	Alterar os valores no arquivo localfile.properties do Trino.	Not available.
trino-connector-memory	Alterar os valores no arquivo memory.properties do Trino.	Not available.
trino-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Trino.	Not available.
trino-connector-mysql	Alterar os valores no arquivo mysql.properties do Trino.	Not available.
trino-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Trino.	Not available.
trino-connector-raptor	Alterar os valores no arquivo raptor.properties do Trino.	Not available.
trino-connector-redis	Alterar os valores no arquivo redis.properties do Trino.	Not available.
trino-connector-redshift	Alterar os valores no arquivo redshift.properties do Trino.	Not available.
trino-connector-tpch	Alterar os valores no arquivo tpch.properties do Trino.	Not available.
trino-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Trino.	Not available.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j2	Alterar os valores no arquivo log4j2.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.

Classificações	Descrição	Ações de reconfiguração
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie and HiveServe r2.
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.

Classificações	Descrição	Ações de reconfiguração
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Log de alterações 6.8.0

Log de alterações da versão 6.8.0 e das notas da versão

Data	Event	Descrição
2023-08-21	Atualizar	Adicionado um problema conhecido com o Hadoop 3.3.3.
2023-07-26	Atualizar	Novos rótulos de versão `2.0.20230612.0` e `2.0.20230628.0` do sistema operacional.
2022-09-06	Implantação concluída	Amazon EMR 6.8 totalmente implantado em todas as regiões compatíveis
2022-09-06	Publicação inicial	Notas da versão 6.8 do Amazon EMR publicadas pela primeira vez
2022-08-31	Lançamento inicial	Versão 6.8 do Amazon EMR lançada em regiões comerciais limitadas

Versão 6.7.0 do Amazon EMR

Versões 6.7.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, Iceberg, JupyterEnterpriseGatewayJupyterHub, Livy, MXNet, Oozie, PhoenixPig, Presto, Spark, SqoopTensorFlow, Tez, Trino, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-6.7.0	emr-6.6.0	emr-6.5.0	emr-6.4.0
AWS SDK for Java	1.12.170	1.12.170	1.12.31	1.12.31
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.2	1.14.2	1.14.0	1.13.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.4	2.4.4	2.4.4	2.4.4

	emr-6.7.0	emr-6.6.0	emr-6.5.0	emr-6.4.0
HCatalog	3.1.3	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.3	3.1.2	3.1.2	3.1.2
Hudi	0.11.0-amzn-0	0.10.1-amzn-0	0.9.0-amzn-1	0.8.0-amzn-0
Hue	4.10.0	4.10.0	4.9.0	4.9.0
Iceberg	0.13.1-amzn-0	0.13.1	0.12.0	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.8.0	1.8.0	1.8.0	1.8.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,272	0,267	0,261	0,254,1
Spark	3.2.1	3.2.0	3.1.2	3.1.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-6.7.0	emr-6.6.0	emr-6.5.0	emr-6.4.0
Trino (PrestoSQL)	378	367	360	359
Zeppelin	0.10.0	0.10.0	0.10.0	0.9.0
ZooKeeper	3.5.7	3.5.7	3.5.7	3.5.7

Notas da versão 6.7.0

As notas da versão a seguir incluem informações para a versão 6.7.0 do Amazon EMR. As alterações são referentes à versão 6.6.0.

Data da versão inicial: 15 de julho de 2022

Novos atributos

- O Amazon EMR agora é compatível com Apache Spark 3.2.1, Apache Hive 3.1.3, HUDI 0.11, PrestoDB 0.272 e Trino 0.378.
- Suporta controles de acesso baseados em IAM Role e Lake Formation com etapas do EMR (Spark, Hive) para Amazon EMR em clusters. EC2
- Ele é compatível com instruções de definição de dados do Apache Spark em clusters habilitados para Apache Ranger. Isso agora inclui suporte para aplicações do Trino lendo e gravando metadados do Apache Hive em clusters habilitados para Apache Ranger. Para obter mais informações, consulte Habilitar governança federada usando Trino e Apache Ranger no Amazon EMR.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Europa (Londres) , Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (Oregon) AWS GovCloud , Oeste dos EUA (Norte da Califórnia) AWS GovCloud , Ásia-Pacífico (Cingapura), Ásia-Pací fico (Sydney)), Israel (Tel Aviv), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			, Europa (Espanha), China (Ningxia), Europa (Zurique)
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Hyderabad), Ásia-Pacífico (Jacarta) Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (EAU), Canadá (Central) , (EUA- AWS GovCloud Leste), AWS GovCloud (Oeste dos EUA), China

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Pequim), China (Ningxia), Israel (Tel Aviv)
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Centro Leste (Emirados Árabes Unidos), Israel (Tel Aviv), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 001.0	4.14.352	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 816.0	4.14.350	21 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 809.0	4.14.349	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 719.0	4.14.348	25 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Europa (Espanha), Europa (Zurique)
2.0.2024 223.0	4.14.336	8 de março de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 131.0	4.14.336	14 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 124.0	4.14.336	7 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 109.0	4.14.334	24 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 218.0	4.14.330	2 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 206.0	4.14.330	22 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 116.0	4.14.328	11 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 101.0	4.14.327	16 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 020.1	4.14.326	7 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 012.1	4.14.326	26 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 926.0	4.14.322	19 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 906.0	4.14.322	4 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 822.0	4.14.322	30 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 808.0	4.14.320	24 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 727.0	4.14.320	14 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 719.0	4.14.320	2 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta) , África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 628.0	4.14.318	12 de julho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 612.0	4.14.314	23 de junho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 504.1	4.14.313	16 de maio de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 418.0	4.14.311	3 de maio de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 404.1	4.14.311	18 de abril de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)
2.0.2023 404.0	4.14.311	10 de abril de 2023	Leste dos EUA (Norte da Virgínia), Europa (Paris)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 320.0	4.14.309	30 de março de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 307.0	4.14.305	15 de março de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 207.0	4.14.304	22 de fevereiro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 119.1	4.14.301	3 de fevereiro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 210.1	4.14.301	22 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 103.3	4.14.296	5 de dezembro de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 004.0	4.14.294	2 de novembro de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 912.1	4.14.291	7 de outubro de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)
2.0.2022 719.0	4.14.287	10 de agosto de 2022	`us-west-1` , `eu-west-3` , `eu-north-1` , `ap-south-1` , `me-south-1`

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 606.1	4.14.281	15 de julho de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

Problemas conhecidos

- Quando a versão 6.5.0, 6.6.0 ou 6.7.0 do Amazon EMR lê as tabelas do Apache Phoenix por meio do shell do Apache Spark, ocorre um `NoSuchMethodError` porque o Amazon EMR usa um `Hbase.compat.version` incorreto. A versão 6.8.0 do Amazon EMR corrige esse problema.

- Ao usar o conector DynamoDB com o Spark nas versões 6.6.0, 6.7.0 e 6.8.0 do Amazon EMR, todas as leituras da tabela retornam um resultado vazio, mesmo que a divisão de entrada faça referência a dados que não estão vazios. Isso ocorre porque o Spark 3.2.0 define `spark.hadoopRDD.ignoreEmptySplits` como `true` por padrão. Como solução alternativa, defina explicitamente `spark.hadoopRDD.ignoreEmptySplits` como `false`. A versão 6.9.0 do Amazon EMR corrige esse problema.
- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
  - O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

  Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.
- Com as versões 5.36.0 e 6.6.0 a 6.9.0 do Amazon EMR, os componentes do serviço `SecretAgent` e `RecordServer` podem sofrer perda de dados de log devido a uma configuração incorreta do padrão de nome de arquivo nas propriedades do Log4j2. A configuração incorreta faz com que os componentes gerem somente um arquivo de log por dia. Quando a estratégia de rotação ocorre, ela substitui o arquivo existente em vez de gerar um novo arquivo de log, conforme esperado. Como solução alternativa, use uma ação de bootstrap para gerar arquivos de log a cada hora e acrescentar um número inteiro de incremento automático no nome do arquivo para lidar com a rotação.

  Para as versões 6.6.0 a 6.9.0 do Amazon EMR, use a seguinte ação de bootstrap ao iniciar um cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

Para o Amazon EMR 5.36.0, use a ação de bootstrap a seguir ao iniciar um cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

- A API `GetClusterSessionCredentials` não é compatível com clusters executados no Amazon EMR 6.7 ou versões inferiores.

Versões de componente 6.7.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.1	SDK do Amazon SageMaker Spark
emr-ddb	4.16.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.2.0	Bibliotecas convenientes para o ecossistema do Hadoop.

Componente	Versão	Descrição
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.6.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.22.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	2.1.0	EMR S3Select Connector
emrfs	2.52.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.14.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.14.2	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	3.2.1-amzn-7	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	3.2.1-amzn-7	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.2.1-amzn-7	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.2.1-amzn-7	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	3.2.1-amzn-7	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.2.1-amzn-7	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.2.1-amzn-7	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	3.2.1-amzn-7	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.

Componente	Versão	Descrição
hadoop-yarn-nodemanager	3.2.1-amzn-7	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	3.2.1-amzn-7	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hadoop-yarn-timeline-server	3.2.1-amzn-7	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	2.4.4-amzn-3	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	2.4.4-amzn-3	Serviço para atender uma ou mais HBase regiões.
hbase-client	2.4.4-amzn-3	HBase cliente de linha de comando.
hbase-rest-server	2.4.4-amzn-3	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.4.4-amzn-3	Serviço que fornece um endpoint Thrift para. HBase
hbase-operator-tools	2.4.4-amzn-3	Ferramenta de reparo para HBase clusters Apache.

Componente	Versão	Descrição
hcatalog-client	3.1.3-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.3-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	3.1.3-amzn-0	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.3-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	3.1.3-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	3.1.3-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.3-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.11.0-amzn-0	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.11.0-amzn-0	Biblioteca de pacotes para executar o Presto com o Hudi.

Componente	Versão	Descrição
hudi-trino	0.11.0-amzn-0	Biblioteca de pacotes para execução do Presto com o Hudi.
hudi-spark	0.11.0-amzn-0	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.10.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
iceberg	0.13.1-amzn-0	Apache Iceberg é um formato de tabela aberta para conjuntos de dados analíticos imensos
jupyterhub	1.4.1	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.1-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mxnet	1.8.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.
nvidia-cuda	11.0.194	Drivers NVIDIA e toolkit CUDA

Componente	Versão	Descrição
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitações de fluxo de trabalho do Oozie.
opencv	4.5.0	Biblioteca de Visão Computacional de Código Aberto.
phoenix-library	5.1.2	As bibliotecas do phoenix para servidor e cliente
phoenix-connectors	5.1.2	Conectores do Apache Phoenix para Spark-3
phoenix-query-server	5.1.2	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.272-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.272-amzn-0	O serviço que executa partes de uma consulta.
presto-client	0.272-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.

Componente	Versão	Descrição
trino-coordinator	378-amzn-0	Serviço que aceita consultas e gerencia a execução de consultas entre trino-workers.
trino-worker	378-amzn-0	O serviço que executa partes de uma consulta.
trino-client	378-amzn-0	Cliente de linha de comando do Trino instalado em mestres de espera de um cluster de alta disponibilidade (HA) em que o servidor do Trino não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.0.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	3.2.1-amzn-0	Os clientes da linha de comando do Spark.
spark-history-server	3.2.1-amzn-0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.2.1-amzn-0	O mecanismo de execução em memória para o YARN.

Componente	Versão	Descrição
spark-yarn-slave	3.2.1-amzn-0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
spark-rapids	22.02.0-amzn-1	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.4.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.41+	O servidor Apache HTTP.
zeppelin-server	0.10.0	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.5.7	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.5.7	ZooKeeper cliente de linha de comando.

Classificações de configuração 6.7.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-6.7.0

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop

Classificações	Descrição	Ações de reconfiguração
		Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.

Classificações	Descrição	Ações de reconfiguração
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.

Classificações	Descrição	Ações de reconfiguração
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classificações	Descrição	Ações de reconfiguração
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.

Classificações	Descrição	Ações de reconfiguração
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.	Not available.
iceberg-defaults	Alterar os valores no arquivo iceberg-defaults.conf do Iceberg.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.

Classificações	Descrição	Ações de reconfiguração
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j	Alterar as configurações de log4j.properties no Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.

Classificações	Descrição	Ações de reconfiguração
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
trino-log	Alterar os valores no arquivo log.properties do Trino.	Restarts Trino-Server (for Trino)
trino-config	Alterar os valores no arquivo config.properties do Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Alterar os valores no arquivo password-authenticator.properties do Trino.	Restarts Trino-Server (for Trino)
trino-env	Alterar os valores no arquivo trino-env.sh do Trino.	Restarts Trino-Server (for Trino)
trino-node	Alterar os valores no arquivo node.properties do Trino.	Not available.
trino-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Trino.	Not available.
trino-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Trino.	Not available.
trino-connector-hive	Alterar os valores no arquivo hive.properties do Trino.	Restarts Trino-Server (for Trino)

Classificações	Descrição	Ações de reconfiguração
trino-connector-iceberg	Alterar os valores no arquivo iceberg.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Alterar os valores no arquivo jmx.properties do Trino.	Not available.
trino-connector-kafka	Alterar os valores no arquivo kafka.properties do Trino.	Not available.
trino-connector-localfile	Alterar os valores no arquivo localfile.properties do Trino.	Not available.
trino-connector-memory	Alterar os valores no arquivo memory.properties do Trino.	Not available.
trino-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Trino.	Not available.
trino-connector-mysql	Alterar os valores no arquivo mysql.properties do Trino.	Not available.
trino-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Trino.	Not available.
trino-connector-raptor	Alterar os valores no arquivo raptor.properties do Trino.	Not available.
trino-connector-redis	Alterar os valores no arquivo redis.properties do Trino.	Not available.
trino-connector-redshift	Alterar os valores no arquivo redshift.properties do Trino.	Not available.
trino-connector-tpch	Alterar os valores no arquivo tpch.properties do Trino.	Not available.
trino-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Trino.	Not available.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.

Classificações	Descrição	Ações de reconfiguração
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie and HiveServe r2.
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.

Classificações	Descrição	Ações de reconfiguração
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Versão 6.6.0 do Amazon EMR

Versões 6.6.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, Iceberg, JupyterEnterpriseGatewayJupyterHub, Livy, MXNet, Oozie, PhoenixPig, Presto, Spark, SqoopTensorFlow, Tez, Trino, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-6.6.0	emr-6.5.0	emr-6.4.0	emr-6.3.1
AWS SDK for Java	1.12.170	1.12.31	1.12.31	1.11.97
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-

	emr-6.6.0	emr-6.5.0	emr-6.4.0	emr-6.3.1
Delta	-	-	-	-
Flink	1.14.2	1.14.0	1.13.1	1.12.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.4	2.4.4	2.4.4	2.2.6
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.10.1-amzn-0	0.9.0-amzn-1	0.8.0-amzn-0	0.7.0-amzn-0
Hue	4.10.0	4.9.0	4.9.0	4.9.0
Iceberg	0.13.1	0.12.0	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.2.2
Livy	0.7.1	0.7.1	0.7.1	0.7.0
MXNet	1.8.0	1.8.0	1.8.0	1.7.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,267	0,261	0,254,1	0,245,1
Spark	3.2.0	3.1.2	3.1.2	3.1.1

	emr-6.6.0	emr-6.5.0	emr-6.4.0	emr-6.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	367	360	359	350
Zeppelin	0.10.0	0.10.0	0.9.0	0.9.0
ZooKeeper	3.5.7	3.5.7	3.5.7	3.4.14

Notas da versão 6.6.0

As notas da versão a seguir incluem informações para a versão 6.6.0 do Amazon EMR. As alterações são referentes à versão 6.5.0.

Data da versão inicial: 9 de maio de 2022

Data da documentação atualizada: 15 de junho de 2022

Novos atributos

- O Amazon EMR 6.6 agora é compatível com Apache Spark 3.2, Apache Spark RAPIDS 22.02, CUDA 11, Apache Hudi 0.10.1, Apache Iceberg 0.13, Trino 0.367 e PrestoDB 0.267.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda) , Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Leste dos EUA (Norte da Virgínia) , Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Europa (Londres), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA) AWS GovCloud , Oeste dos EUA (Oregon) AWS GovCloud , Oeste dos EUA (Norte da Califórnia), Ásia-Pacífico (Cingapur a), Ásia-Pacífico (Sydney), Canadá (Central), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 001.0	4.14.352	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 816.0	4.14.350	21 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 809.0	4.14.349	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 719.0	4.14.348	25 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 223.0	4.14.336	8 de março de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 131.0	4.14.336	14 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 124.0	4.14.336	7 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 109.0	4.14.334	24 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 218.0	4.14.330	2 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 206.0	4.14.330	22 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 116.0	4.14.328	11 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 101.0	4.14.327	16 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 020.1	4.14.326	7 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 012.1	4.14.326	26 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 926.0	4.14.322	19 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 906.0	4.14.322	4 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 822.0	4.14.322	30 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 808.0	4.14.320	24 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 727.0	4.14.320	14 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 719.0	4.14.320	2 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta) , África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 628.0	4.14.318	12 de julho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 612.0	4.14.314	23 de junho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 504.1	4.14.313	16 de maio de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 418.0	4.14.311	3 de maio de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 404.1	4.14.311	18 de abril de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)
2.0.2023 404.0	4.14.311	10 de abril de 2023	Leste dos EUA (Norte da Virgínia), Europa (Paris)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 320.0	4.14.309	30 de março de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 307.0	4.14.305	15 de março de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 207.0	4.14.304	22 de fevereiro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 119.1	4.14.301	3 de fevereiro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 210.1	4.14.301	22 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 103.3	4.14.296	5 de dezembro de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 004.0	4.14.294	2 de novembro de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 912.1	4.14.291	7 de outubro de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)
2.0.2022 805.0	4.14.287	30 de agosto de 2022	`us-west-1`

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 719.0	4.14.287	10 de agosto de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 426.0	4.14.281	10 de junho de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 406.1	4.14.275	2 de maio de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

- Com as versões 6.6 e posteriores do Amazon EMR,as aplicações que usam o Log4j 1.x e o Log4j 2.x são atualizadas para usar o Log4j 1.2.17 (ou superior) e o Log4j 2.17.1 (ou superior), respectivamente, e não exigem o uso de ações de bootstrap para mitigar os problemas de CVE.
- [Ajuste de escala gerenciado] Otimização do ajuste de escala gerenciado de dados embaralhados do Spark: para as versões 5.34.0 e posteriores do Amazon EMR e as versões 6.4.0 e posteriores do EMR, o ajuste de escala gerenciado agora reconhece dados embaralhados do Spark (dados que o Spark redistribui entre partições para executar operações específicas). Para obter mais

informações sobre operações de shuffle, consulte Usar ajuste de escala gerenciado do EMR no Amazon EMR no Guia de gerenciamento do Amazon EMR e no Guia de programação do Spark.

- Desde as versões 5.32.0 e 6.5.0 do Amazon EMR, o dimensionamento do executor dinâmico para o Apache Spark está habilitado por padrão. Para ativar ou desativar esse atributo, você pode usar o parâmetro de configuração `spark.yarn.heterogeneousExecutors.enabled`.

Alterações, melhorias e problemas resolvidos

- O Amazon EMR reduz o tempo de inicialização do cluster em até 80 segundos, em média, para clusters que usam a opção de AMI padrão do EMR e só instalam aplicações comuns, como Apache Hadoop, Apache Spark e Apache Hive.

Problemas conhecidos

- Quando a versão 6.5.0, 6.6.0 ou 6.7.0 do Amazon EMR lê as tabelas do Apache Phoenix por meio do shell do Apache Spark, ocorre um `NoSuchMethodError` porque o Amazon EMR usa um `Hbase.compat.version` incorreto. A versão 6.8.0 do Amazon EMR corrige esse problema.
- Ao usar o conector DynamoDB com o Spark nas versões 6.6.0, 6.7.0 e 6.8.0 do Amazon EMR, todas as leituras da tabela retornam um resultado vazio, mesmo que a divisão de entrada faça referência a dados que não estão vazios. Isso ocorre porque o Spark 3.2.0 define `spark.hadoopRDD.ignoreEmptySplits` como `true` por padrão. Como solução alternativa, defina explicitamente `spark.hadoopRDD.ignoreEmptySplits` como `false`. A versão 6.9.0 do Amazon EMR corrige esse problema.
- Em clusters de execução prolongada do Trino, o Amazon EMR 6.6.0 habilita os parâmetros de registro em log da coleta de resíduos no jvm.config do Trino para obter melhores insights dos registros em log da coleta de resíduos. Essa alteração anexa muitos registros da coleta de lixo ao arquivo launcher.log (/var/log/trino/launcher.log). Se você estiver executando clusters do Trino no Amazon EMR 6.6.0, poderá encontrar nós que perderão espaço em disco depois que o cluster estiver em execução por alguns dias devido aos logs anexados.

  A solução alternativa para esse problema é executar o script abaixo como uma ação de bootstrap para desativar os parâmetros de registro em log da coleta de resíduos no jvm.config ao criar ou clonar o cluster para o Amazon EMR 6.6.0.

```
#!/bin/bash
  set -ex
  PRESTO_PUPPET_DIR='/var/aws/emr/bigtop-deploy/puppet/modules/trino'
```

```
  sudo bash -c "sed -i '/-Xlog/d' ${PRESTO_PUPPET_DIR}/templates/jvm.config"
```

- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
  - O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

  Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.
- Com as versões 5.36.0 e 6.6.0 a 6.9.0 do Amazon EMR, os componentes do serviço `SecretAgent` e `RecordServer` podem sofrer perda de dados de log devido a uma configuração incorreta do padrão de nome de arquivo nas propriedades do Log4j2. A configuração incorreta faz com que os componentes gerem somente um arquivo de log por dia. Quando a estratégia de rotação ocorre, ela substitui o arquivo existente em vez de gerar um novo arquivo de log, conforme esperado. Como solução alternativa, use uma ação de bootstrap para gerar arquivos de log a cada hora e acrescentar um número inteiro de incremento automático no nome do arquivo para lidar com a rotação.

  Para as versões 6.6.0 a 6.9.0 do Amazon EMR, use a seguinte ação de bootstrap ao iniciar um cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

  Para o Amazon EMR 5.36.0, use a ação de bootstrap a seguir ao iniciar um cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

Versões de componente 6.6.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.1	SDK do Amazon SageMaker Spark
emr-ddb	4.16.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.2.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.5.0	Ambiente Conda para bloco de anotações do emr, que

Componente	Versão	Descrição
		inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.20.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	2.1.0	EMR S3Select Connector
emrfs	2.50.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.14.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.14.2	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.

Componente	Versão	Descrição
hadoop-client	3.2.1-amzn-6	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	3.2.1-amzn-6	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.2.1-amzn-6	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.2.1-amzn-6	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	3.2.1-amzn-6	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.2.1-amzn-6	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.2.1-amzn-6	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	3.2.1-amzn-6	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.2.1-amzn-6	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	3.2.1-amzn-6	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	3.2.1-amzn-6	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	2.4.4-amzn-2	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	2.4.4-amzn-2	Serviço para atender uma ou mais HBase regiões.
hbase-client	2.4.4-amzn-2	HBase cliente de linha de comando.
hbase-rest-server	2.4.4-amzn-2	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.4.4-amzn-2	Serviço que fornece um endpoint Thrift para. HBase
hbase-operator-tools	2.4.4-amzn-2	Ferramenta de reparo para HBase clusters Apache.
hcatalog-client	3.1.2-amzn-7	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.

Componente	Versão	Descrição
hcatalog-server	3.1.2-amzn-7	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	3.1.2-amzn-7	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.2-amzn-7	O cliente da linha de comando do Hive.
hive-hbase	3.1.2-amzn-7	Cliente do Hive-hbase.
hive-metastore-server	3.1.2-amzn-7	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.2-amzn-7	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.10.1-amzn-0	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.10.1-amzn-0	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-trino	0.10.1-amzn-0	Biblioteca de pacotes para execução do Presto com o Hudi.

Componente	Versão	Descrição
hudi-spark	0.10.1-amzn-0	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.10.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
iceberg	0.13.1	Apache Iceberg é um formato de tabela aberta para conjuntos de dados analíticos imensos
jupyterhub	1.4.1	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.1-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mxnet	1.8.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.
nvidia-cuda	11.0.194	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.

Componente	Versão	Descrição
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.5.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	5.1.2	As bibliotecas do phoenix para servidor e cliente
phoenix-connectors	5.1.2	Conectores do Apache Phoenix para Spark-3
phoenix-query-server	5.1.2	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.267-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.267-amzn-0	O serviço que executa partes de uma consulta.
presto-client	0.267-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
trino-coordinator	367-amzn-0	Serviço que aceita consultas e gerencia a execução de consultas entre trino-workers.

Componente	Versão	Descrição
trino-worker	367-amzn-0	O serviço que executa partes de uma consulta.
trino-client	367-amzn-0	Cliente de linha de comando do Trino instalado em mestres de espera de um cluster de alta disponibilidade (HA) em que o servidor do Trino não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.0.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	3.2.0-amzn-0	Os clientes da linha de comando do Spark.
spark-history-server	3.2.0-amzn-0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.2.0-amzn-0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.2.0-amzn-0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.

Componente	Versão	Descrição
spark-rapids	22.02.0-amzn-0	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.4.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.41+	O servidor Apache HTTP.
zeppelin-server	0.10.0	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.5.7	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.5.7	ZooKeeper cliente de linha de comando.

Classificações de configuração 6.6.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-6.6.0

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.

Classificações	Descrição	Ações de reconfiguração
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.

Classificações	Descrição	Ações de reconfiguração
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.

Classificações	Descrição	Ações de reconfiguração
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classificações	Descrição	Ações de reconfiguração
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.

Classificações	Descrição	Ações de reconfiguração
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.	Not available.
iceberg-defaults	Alterar os valores no arquivo iceberg-defaults.conf do Iceberg.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.

Classificações	Descrição	Ações de reconfiguração
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j	Alterar as configurações de log4j.properties no Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.

Classificações	Descrição	Ações de reconfiguração
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
trino-log	Alterar os valores no arquivo log.properties do Trino.	Restarts Trino-Server (for Trino)
trino-config	Alterar os valores no arquivo config.properties do Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Alterar os valores no arquivo password-authenticator.properties do Trino.	Restarts Trino-Server (for Trino)
trino-env	Alterar os valores no arquivo trino-env.sh do Trino.	Restarts Trino-Server (for Trino)
trino-node	Alterar os valores no arquivo node.properties do Trino.	Not available.
trino-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Trino.	Not available.
trino-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Trino.	Not available.
trino-connector-hive	Alterar os valores no arquivo hive.properties do Trino.	Restarts Trino-Server (for Trino)

Classificações	Descrição	Ações de reconfiguração
trino-connector-iceberg	Alterar os valores no arquivo iceberg.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Alterar os valores no arquivo jmx.properties do Trino.	Not available.
trino-connector-kafka	Alterar os valores no arquivo kafka.properties do Trino.	Not available.
trino-connector-localfile	Alterar os valores no arquivo localfile.properties do Trino.	Not available.
trino-connector-memory	Alterar os valores no arquivo memory.properties do Trino.	Not available.
trino-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Trino.	Not available.
trino-connector-mysql	Alterar os valores no arquivo mysql.properties do Trino.	Not available.
trino-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Trino.	Not available.
trino-connector-raptor	Alterar os valores no arquivo raptor.properties do Trino.	Not available.
trino-connector-redis	Alterar os valores no arquivo redis.properties do Trino.	Not available.
trino-connector-redshift	Alterar os valores no arquivo redshift.properties do Trino.	Not available.
trino-connector-tpch	Alterar os valores no arquivo tpch.properties do Trino.	Not available.
trino-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Trino.	Not available.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.

Classificações	Descrição	Ações de reconfiguração
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie and HiveServe r2.
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.

Classificações	Descrição	Ações de reconfiguração
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Versão 6.5.0 do Amazon EMR

Versões 6.5.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, Iceberg, JupyterEnterpriseGatewayJupyterHub, Livy, MXNet, Oozie, PhoenixPig, Presto, Spark, SqoopTensorFlow, Tez, Trino, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-6.5.0	emr-6.4.0	emr-6.3.1	emr-6.3.0
AWS SDK for Java	1.12.31	1.12.31	1.11.977	1.11.977
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-

	emr-6.5.0	emr-6.4.0	emr-6.3.1	emr-6.3.0
Delta	-	-	-	-
Flink	1.14.0	1.13.1	1.12.1	1.12.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.4	2.4.4	2.2.6	2.2.6
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.9.0-amzn-1	0.8.0-amzn-0	0.7.0-amzn-0	0.7.0-amzn-0
Hue	4.9.0	4.9.0	4.9.0	4.9.0
Iceberg	0.12.0	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.2.2	1.2.2
Livy	0.7.1	0.7.1	0.7.0	0.7.0
MXNet	1.8.0	1.8.0	1.7.0	1.7.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,261	0,254,1	0,245,1	0,245,1
Spark	3.1.2	3.1.2	3.1.1	3.1.1

	emr-6.5.0	emr-6.4.0	emr-6.3.1	emr-6.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	360	359	350	350
Zeppelin	0.10.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.5.7	3.5.7	3.4.14	3.4.14

Notas da versão 6.5.0

As notas da versão a seguir incluem informações para a versão 6.5.0 do Amazon EMR. As alterações são referentes à versão 6.4.0.

Data da versão inicial: 20 de janeiro de 2022

Data da versão atualizada: 21 de março de 2022

Novos atributos

- [Ajuste de escala gerenciado] Otimização do ajuste de escala gerenciado de dados embaralhados do Spark: para as versões 5.34.0 e posteriores do Amazon EMR e as versões 6.4.0 e posteriores do EMR, o ajuste de escala gerenciado agora reconhece dados embaralhados do Spark (dados que o Spark redistribui entre partições para executar operações específicas). Para obter mais informações sobre operações de shuffle, consulte Usar ajuste de escala gerenciado do EMR no Amazon EMR no Guia de gerenciamento do Amazon EMR e no Guia de programação do Spark.
- Desde as versões 5.32.0 e 6.5.0 do Amazon EMR, o dimensionamento do executor dinâmico para o Apache Spark está habilitado por padrão. Para ativar ou desativar esse atributo, você pode usar o parâmetro de configuração `spark.yarn.heterogeneousExecutors.enabled`.
- Suporte para o formato de tabela aberta Apache Iceberg para conjuntos de dados analíticos imensos.
- Support para ranger-trino-plugin 2.0.1-amzn-1

- Suporte para toree 0.5.0

Alterações, melhorias e problemas resolvidos

- A versão de lançamento do Amazon EMR 6.5 agora é compatível com o Apache Iceberg 0.12.0 e fornece melhorias no runtime com o Ambiente de Tempo de Execução do Amazon EMR para Apache Spark, o Ambiente de Tempo de Execução do Amazon EMR para Presto e o Runtime do Amazon EMR para Apache Hive.
- O Apache Iceberg é um formato de tabela aberta para grandes conjuntos de dados no Amazon S3 que fornece performance rápida de consultas em tabelas grandes, confirmações atômicas, gravações simultâneas e evolução de tabelas compatível com SQL. Com o EMR 6.5, você pode usar o Apache Spark 3.1.2 com o formato de tabela Iceberg.
- O Apache Hudi 0.9 adiciona suporte a DDL e DML do Spark SQL. Isso permite a criação e a atualização de tabelas do Hudi usando apenas instruções de SQL. O Apache Hudi 0.9 também inclui melhorias na performance do lado da consulta e do lado do gravador.
- O Runtime do Amazon EMR para Apache Hive melhora a performance do Apache Hive no Amazon S3 ao remover operações de renomeação durante operações de preparação e melhora a performance dos comandos de verificação do metastore (MSCK) usados para reparar tabelas.

Problemas conhecidos

- Quando a versão 6.5.0, 6.6.0 ou 6.7.0 do Amazon EMR lê as tabelas do Apache Phoenix por meio do shell do Apache Spark, ocorre um `NoSuchMethodError` porque o Amazon EMR usa um `Hbase.compat.version` incorreto. A versão 6.8.0 do Amazon EMR corrige esse problema.
- Os clusters do pacote do Hbase em alta disponibilidade (HA) apresentam falha no provisionamento com o tamanho de volume e o tipo de instância padrão. A solução alternativa para esse problema é aumentar o tamanho do volume raiz.
- Para usar as ações do Spark com o Apache Oozie, você deve adicionar a seguinte configuração ao seu arquivo `workflow.xml` do Oozie. Caso contrário, várias bibliotecas críticas, como Hadoop e EMRFS, estarão ausentes do classpath dos executores do Spark que o Oozie inicia.

```
<spark-opts>--conf spark.yarn.populateHadoopClasspath=true</spark-opts>
```

- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR,

pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:

- Duas ou mais partições são verificadas na mesma tabela.
- Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
- O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.

Versões de componente 6.5.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.1	SDK do Amazon SageMaker Spark

Componente	Versão	Descrição
emr-ddb	4.16.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.2.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.4.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.19.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	2.1.0	EMR S3Select Connector
emrfs	2.48.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.14.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.14.0	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager

Componente	Versão	Descrição
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	3.2.1-amzn-5	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	3.2.1-amzn-5	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.2.1-amzn-5	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.2.1-amzn-5	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	3.2.1-amzn-5	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.

Componente	Versão	Descrição
hadoop-httpfs-server	3.2.1-amzn-5	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.2.1-amzn-5	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-5	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.2.1-amzn-5	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	3.2.1-amzn-5	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	3.2.1-amzn-5	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	2.4.4-amzn-1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	2.4.4-amzn-1	Serviço para atender uma ou mais HBase regiões.
hbase-client	2.4.4-amzn-1	HBase cliente de linha de comando.

Componente	Versão	Descrição
hbase-rest-server	2.4.4-amzn-1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.4.4-amzn-1	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	3.1.2-amzn-6	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.2-amzn-6	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	3.1.2-amzn-6	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.2-amzn-6	O cliente da linha de comando do Hive.
hive-hbase	3.1.2-amzn-6	Cliente do Hive-hbase.
hive-metastore-server	3.1.2-amzn-6	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.2-amzn-6	O serviço que aceita as consultas do Hive como solicitações da web.

Componente	Versão	Descrição
hudi	0.9.0-amzn-1	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.9.0-amzn-1	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-trino	0.9.0-amzn-1	Biblioteca de pacotes para execução do Presto com o Hudi.
hudi-spark	0.9.0-amzn-1	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.9.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
iceberg	0.12.0	Apache Iceberg é um formato de tabela aberta para conjuntos de dados analíticos imensos
jupyterhub	1.4.1	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.1-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP

Componente	Versão	Descrição
mxnet	1.8.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.
nvidia-cuda	10.1.243	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.5.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	5.1.2	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	5.1.2	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.261-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.261-amzn-0	O serviço que executa partes de uma consulta.

Componente	Versão	Descrição
presto-client	0.261-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
trino-coordinator	360	Serviço que aceita consultas e gerencia a execução de consultas entre trino-workers.
trino-worker	360	O serviço que executa partes de uma consulta.
trino-client	360	Cliente de linha de comando do Trino instalado em mestres de espera de um cluster de alta disponibilidade (HA) em que o servidor do Trino não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.0.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	3.1.2-amzn-1	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	3.1.2-amzn-1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.1.2-amzn-1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.1.2-amzn-1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
spark-rapids	0.4.1	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.4.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.41+	O servidor Apache HTTP.
zeppelin-server	0.10.0	O notebook baseado na web que permite um data analytics interativo.

Componente	Versão	Descrição
zookeeper-server	3.5.7	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.5.7	ZooKeeper cliente de linha de comando.

Classificações de configuração 6.5.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-6.5.0

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Classificações	Descrição	Ações de reconfiguração
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.

Classificações	Descrição	Ações de reconfiguração
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.

Classificações	Descrição	Ações de reconfiguração
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.	Not available.
iceberg-defaults	Alterar os valores no arquivo iceberg-defaults.conf do Iceberg.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j	Alterar as configurações de log4j.properties no Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)

Classificações	Descrição	Ações de reconfiguração
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
trino-log	Alterar os valores no arquivo log.properties do Trino.	Restarts Trino-Server (for Trino)
trino-config	Alterar os valores no arquivo config.properties do Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Alterar os valores no arquivo password-authenticator.properties do Trino.	Restarts Trino-Server (for Trino)

Classificações	Descrição	Ações de reconfiguração
trino-env	Alterar os valores no arquivo trino-env.sh do Trino.	Restarts Trino-Server (for Trino)
trino-node	Alterar os valores no arquivo node.properties do Trino.	Not available.
trino-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Trino.	Not available.
trino-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Trino.	Not available.
trino-connector-hive	Alterar os valores no arquivo hive.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Alterar os valores no arquivo jmx.properties do Trino.	Not available.
trino-connector-kafka	Alterar os valores no arquivo kafka.properties do Trino.	Not available.
trino-connector-localfile	Alterar os valores no arquivo localfile.properties do Trino.	Not available.
trino-connector-memory	Alterar os valores no arquivo memory.properties do Trino.	Not available.
trino-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Trino.	Not available.
trino-connector-mysql	Alterar os valores no arquivo mysql.properties do Trino.	Not available.
trino-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Trino.	Not available.
trino-connector-raptor	Alterar os valores no arquivo raptor.properties do Trino.	Not available.

Classificações	Descrição	Ações de reconfiguração
trino-connector-redis	Alterar os valores no arquivo redis.properties do Trino.	Not available.
trino-connector-redshift	Alterar os valores no arquivo redshift.properties do Trino.	Not available.
trino-connector-tpch	Alterar os valores no arquivo tpch.properties do Trino.	Not available.
trino-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Trino.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.

Classificações	Descrição	Ações de reconfiguração
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie and HiveServe r2.
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Versão 6.4.0 do Amazon EMR

Versões 6.4.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterEnterpriseGateway, JupyterHubLivy, MXNet, Oozie, Phoenix, PigPresto, Spark, Sqoop, TensorFlowTez, Trino, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR

- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-6.4.0	emr-6.3.1	emr-6.3.0	emr-6.2.1
AWS SDK for Java	1.12.31	1.11.977	1.11.977	1.11.880
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.13.1	1.12.1	1.12.1	1.11.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.4	2.2.6	2.2.6	2.2.6-amzn-0
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.8.0-amzn-0	0.7.0-amzn-0	0.7.0-amzn-0	0.6.0-amzn-1
Hue	4.9.0	4.9.0	4.9.0	4.8.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0

	emr-6.4.0	emr-6.3.1	emr-6.3.0	emr-6.2.1
JupyterHub	1.4.1	1.2.2	1.2.2	1.1.0
Livy	0.7.1	0.7.0	0.7.0	0.7.0
MXNet	1.8.0	1.7.0	1.7.0	1.7.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.0
Phoenix	5.1.2	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,254,1	0,245,1	0,245,1	0,238,3
Spark	3.1.2	3.1.1	3.1.1	3.0.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.3.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	359	350	350	343
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.5.7	3.4.14	3.4.14	3.4.14

Notas da versão 6.4.0

As notas da versão a seguir incluem informações para a versão 6.4.0 do Amazon EMR. As alterações são referentes à versão 6.3.0.

Data da versão inicial: 20 de setembro de 2021

Data da versão atualizada: 21 de março de 2022

Aplicações compatíveis

- AWS SDK para Java versão 1.12.31
- CloudWatch Sink versão 2.2.0
- Conector do DynamoDB versão 4.16.0
- EMRFS versão 2.47.0
- Amazon EMR Goodies versão 3.2.0
- Conector Kinesis versão 3.5.0 do Amazon EMR
- Servidor de registros do Amazon EMR versão 2.1.0
- Scripts versão 2.5.0 do Amazon EMR
- Flink versão 1.13.1
- Ganglia versão 3.7.2
- AWS Glue Hive Metastore Client versão 3.3.0
- Hadoop versão 3.2.1-amzn-4
- HBase versão 2.4.4-amzn-0
- HBase-operator-tools 1.1.0
- HCatalog versão 3.1.2-amzn-5
- Hive versão 3.1.2-amzn-5
- Hudi versão 0.8.0-amzn-0
- Hue versão 4.9.0
- Java JDK versão Corretto-8.302.08.1 (compilação 1.8.0_302-b08)
- JupyterHub versão 1.4.1
- Livy versão 0.7.1-incubating
- MXNet versão 1.8.0
- Oozie versão 5.2.1
- Phoenix versão 5.1.2
- Pig versão 0.17.0
- Presto versão 0.254.1-amzn-0
- Trino versão 359
- KMS do Apache Ranger (criptografia transparente multi-mestre) versão 2.0.0

- ranger-plugins 2.0.1-amzn-0
- ranger-s3-plugin 1.2.0
- SageMaker SDK do Spark versão 1.4.1
- Scala versão 2.12.10 (VM de servidor OpenJDK de 64 bits, Java 1.8.0_282)
- Spark versão 3.1.2-amzn-0
- spark-rapids 0.4.1
- Sqoop versão 1.4.7
- TensorFlow versão 2.4.1
- Tez versão 0.9.2
- Zeppelin versão 0.9.0
- Zookeeper versão 3.5.7
- Conectores e drivers: DynamoDB Connector 4.16.0

Novos recursos

- [Ajuste de escala gerenciado] Otimização do ajuste de escala gerenciado de dados embaralhados do Spark: para as versões 5.34.0 e posteriores do Amazon EMR e as versões 6.4.0 e posteriores do EMR, o ajuste de escala gerenciado agora reconhece dados embaralhados do Spark (dados que o Spark redistribui entre partições para executar operações específicas). Para obter mais informações sobre operações de shuffle, consulte Usar ajuste de escala gerenciado do EMR no Amazon EMR no Guia de gerenciamento do Amazon EMR e no Guia de programação do Spark.
- Em clusters do Amazon EMR habilitados para Apache Ranger, você pode usar o Apache Spark SQL para inserir dados ou atualizar as tabelas de metastore do Apache Hive usando `INSERT INTO`, `INSERT OVERWRITE` e `ALTER TABLE`. Ao ser usado ALTER TABLE com Spark SQL, o local da partição deve ser o diretório filho do local de uma tabela. Atualmente,o Amazon EMR não permite a inserção de dados em uma partição cuja localização seja diferente da localização da tabela.
- O PrestoSQL foi renomeado como Trino.
- Hive: a execução de consultas SELECT simples com a cláusula LIMIT é acelerada com a interrupção da execução da consulta assim que o número de registros mencionados na cláusula LIMIT é obtido. As consultas SELECT simples são consultas que não têm a cláusula GROUP BY/ORDER by ou consultas que não têm um estágio redutor. Por exemplo, `.SELECT * from <TABLE> WHERE <Condition> LIMIT <Number>`

Controle de simultaneidade do Hudi

- O Hudi, agora, é compatível com o Optimistic Concurrency Control (OCC - Controle de simultaneidade otimista), que pode ser aproveitado com operações de gravação como UPSERT e INSERT para permitir alterações de vários gravadores na mesma tabela do Hudi. Esse é o OCC em nível de arquivo. Portanto, quaisquer duas confirmações (ou gravadores) podem gravar na mesma tabela, se suas alterações não são conflitantes. Para obter mais informações, consulte o Controle de simultaneidade do Hudi.
- Os clusters do Amazon EMR têm o Zookeeper instalado, que pode ser usado como provedor de bloqueio para o OCC. Para facilitar o uso desse atributo, os clusters do Amazon EMR têm as seguintes propriedades pré-configuradas:

```
hoodie.write.lock.provider=org.apache.hudi.client.transaction.lock.ZookeeperBasedLockProvider
hoodie.write.lock.zookeeper.url=<EMR Zookeeper URL>
hoodie.write.lock.zookeeper.port=<EMR Zookeeper Port>
hoodie.write.lock.zookeeper.base_path=/hudi
```

  Para habilitar o OCC, você precisa configurar as seguintes propriedades com as opções de trabalho do Hudi ou em nível de cluster usando a API de configurações do Amazon EMR:

```
hoodie.write.concurrency.mode=optimistic_concurrency_control
hoodie.cleaner.policy.failed.writes=LAZY (Performs cleaning of failed writes lazily
 instead of inline with every write)
hoodie.write.lock.zookeeper.lock_key=<Key to uniquely identify the Hudi table> (Table
 Name is a good option)
```

Monitoramento Hudi: CloudWatch integração com a Amazon para reportar Hudi Metrics

- O Amazon EMR oferece suporte à publicação de Hudi Metrics na Amazon. CloudWatch Isso é habilitado com a definição das seguintes configurações necessárias:

```
hoodie.metrics.on=true
hoodie.metrics.reporter.type=CLOUDWATCH
```

- A seguir, são mostradas as configurações opcionais do Hudi que você pode alterar:

Configuração	Descrição	Valor
hoodie.metrics.cloudwatch.report.period.seconds	Frequência (em segundos) na qual reportar métricas para a Amazon CloudWatch	O valor padrão é 60s, o que é adequado para a resolução padrão de um minuto oferecida pela Amazon CloudWatch
hoodie.metrics.cloudwatch.metric.prefix	Prefixo a ser adicionado ao nome de cada métrica	O valor padrão está vazio (sem prefixo)
hoodie.metrics.cloudwatch.namespace	CloudWatch Namespace da Amazon sob o qual as métricas são publicadas	O valor padrão é Hudi
hoodie.metrics.cloudwatch.maxDatumsPerSolicitação	Número máximo de dados a serem incluídos em uma solicitação à Amazon CloudWatch	O valor padrão é 20, que é o mesmo que o CloudWatch padrão da Amazon

Suporte e melhorias das configurações do Hudi no Amazon EMR

- Agora, os clientes podem aproveitar a API de configurações e o atributo de reconfiguração do EMR para definir as configurações do Hudi em nível de cluster. Um novo suporte de configuração baseado em arquivos foi introduzido via/etc/hudi/conf/hudi-defaults.conf nos moldes de outros aplicativos como Spark, Hive etc. O EMR configura alguns padrões para melhorar a experiência do usuário:

  — `hoodie.datasource.hive_sync.jdbcurl` está configurado para a URL do servidor do Hive do cluster e não precisa mais ser especificado. Isso é particularmente útil na execução de um trabalho no modo de cluster do Spark, em que anteriormente era necessário especificar o IP principal do Amazon EMR.

  — configurações HBase específicas, que são úteis para usar o HBase índice com o Hudi.

  — Configuração específica do provedor de bloqueio do Zookeeper, conforme discutido em Controle de simultaneidade, o que facilita o uso do Controle de simultaneidade otimista (OCC).

- Alterações adicionais foram introduzidas para reduzir o número de configurações que você precisa passar e inferir automaticamente sempre que possível:

  — A palavra-chave `partitionBy` pode ser usada para especificar a coluna de partição.

  — Ao habilitar o Hive Sync, não é mais obrigatório passar `HIVE_TABLE_OPT_KEY`, `HIVE_PARTITION_FIELDS_OPT_KEY`, `HIVE_PARTITION_EXTRACTOR_CLASS_OPT_KEY`. Esses valores podem ser deduzidos com base no nome da tabela Hudi e no campo de partição.

  — não é obrigatório passar `KEYGENERATOR_CLASS_OPT_KEY`, que pode ser inferido com base em casos mais simples de `SimpleKeyGenerator` e `ComplexKeyGenerator`.

Advertências do Hudi

- O Hudi não permite execução vetorizada no Hive de tabelas Merge on Read (MoR - Mesclar na leitura) e Bootstrap. Por exemplo, `count(*)` apresenta falha com a tabela do Hudi em tempo real quando `hive.vectorized.execution.enabled` está definido como verdadeiro. Como solução alternativa, você pode desabilitar a leitura vetorizada configurando `hive.vectorized.execution.enabled` como `false`.
- O suporte a vários gravadores não é compatível com o atributo de bootstrap do Hudi.
- O streamer e o SQL do Flink são atributos experimentais nesta versão. Esses atributos não são recomendados para uso em implantações de produção.

Alterações, melhorias e problemas resolvidos

Esta é uma versão para corrigir problemas com o Amazon EMR Scaling quando ele não consegue up/scale reduzir a escala de um cluster com sucesso ou causa falhas no aplicativo.

- Anteriormente, a reinicialização manual do gerenciador de recursos em um cluster multimestre fazia com que os daemons do Amazon EMR no cluster, como o Zookeeper, recarregassem todos os nós anteriormente desativados ou perdidos no arquivo znode do Zookeeper. Isso fez com que os limites padrão fossem excedidos em determinadas situações. O Amazon EMR agora remove os registros de nós desativados ou perdidos há mais de uma hora do arquivo do Zookeeper e os limites internos foram aumentados.
- Corrigido um problema em que as solicitações de escalabilidade falhavam em um cluster grande e altamente utilizado quando os daemons do Amazon EMR no cluster estavam executando atividades de verificação de integridade, como a coleta do estado do nó do YARN e o estado do nó

do HDFS. Isso estava acontecendo porque os daemons no cluster não conseguiam comunicar os dados do status de integridade de um nó aos componentes internos do Amazon EMR.

- Aprimorados os daemons do EMR no cluster para rastrear corretamente os estados dos nós quando são reutilizados endereços IP para melhorar a confiabilidade durante operações de escalabilidade.
- SPARK-29683. Corrigido um problema em que ocorriam falhas de trabalho durante a redução da escala verticalmente do cluster, pois o Spark presumia que todos os nós disponíveis estavam na lista de negação.
- YARN-9011. Corrigido um problema em que ocorriam falhas de trabalho devido a uma condição de corrida na desativação do YARN quando o cluster tentava aumentar ou reduzir a escala verticalmente.
- Corrigido problema com falhas de etapas ou tarefas durante a escalabilidade do cluster ao ser garantido que os estados dos nós fossem sempre consistentes entre os daemons do Amazon EMR no cluster e o YARN/HDFS.
- Corrigido um problema em que operações de cluster, como redução de escala verticalmente e envio de etapas, falhavam para clusters do Amazon EMR habilitados com a autenticação Kerberos. Isso ocorreu porque o daemon no cluster do Amazon EMR não renovou o tíquete Kerberos, que é necessário para se comunicar com segurança com a execução no nó primário. HDFS/YARN
- Configurar um cluster para corrigir problemas de desempenho do servidor de linha do tempo do Apache YARN versões 1 e 1.5

  As versões 1 e 1.5 do servidor de linha do tempo do Apache YARN podem causar problemas de performance com clusters do EMR muito ativos e grandes, especialmente com `yarn.resourcemanager.system-metrics-publisher.enabled=true`, que é a configuração padrão no Amazon EMR. Um servidor de linha do tempo do YARN v2 de código aberto resolve o problema de performance relacionado à escalabilidade do servidor de linha do tempo do YARN.

  Outras soluções alternativas para esse problema incluem:

  - Configurando yarn.resourcemanager. system-metrics-publisher.enabled=false em yarn-site.xml.
  - Habilitar a correção para esse problema na criação de um cluster, conforme descrito abaixo.

  As seguintes versões do Amazon EMR contêm uma correção para esse problema de performance do servidor de linha do tempo do YARN.

EMR 5.30.2, 5.31.1, 5.32.1, 5.33.1, 5.34.x, 6.0.1, 6.1.1, 6.2.1, 6.3.1, 6.4.x

Para habilitar a correção em qualquer uma das versões do Amazon EMR especificadas acima, defina essas propriedades como `true` em um arquivo JSON de configurações que é passado usando o parâmetro de comando `aws emr create-cluster`: `--configurations file://./configurations.json`. Ou habilite a correção usando a interface do usuário do console de reconfiguração.

Exemplo de conteúdo do arquivo configurations.json:

```
[
{
"Classification": "yarn-site",
"Properties": {
"yarn.resourcemanager.system-metrics-publisher.timeline-server-v1.enable-batch":
 "true",
"yarn.resourcemanager.system-metrics-publisher.enabled": "true"
},
"Configurations": []
}
]
```

- WebHDFS e o servidor HttpFS estão desabilitados por padrão. Você pode reabilitar o WebHDFS usando a configuração do Hadoop, `dfs.webhdfs.enabled`. O servidor HttpFS pode ser iniciado usando `sudo systemctl start hadoop-httpfs`.
- O HTTPS agora está habilitado por padrão para repositórios do Amazon Linux. Se você estiver usando uma política de VPCE do Amazon S3 para restringir o acesso a buckets específicos, deverá adicionar o novo ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` do bucket do Amazon Linux à sua política (substitua `$region` pela região em que o endpoint está situado). Para obter mais informações, consulte esse tópico nos fóruns de AWS discussão. Anúncio: o Amazon Linux 2 agora oferece suporte à capacidade de usar HTTPS ao se conectar a repositórios de pacotes .
- Hive: a performance da consulta de gravação foi aprimorada ao ser permitido o uso de um diretório temporário no HDFS para o último trabalho. Os dados temporários do trabalho final são gravados no HDFS em vez de no Amazon S3 e a performance é melhorada porque os dados são movidos do HDFS para a localização da tabela final (Amazon S3) em vez de entre dispositivos do Amazon S3.

- Hive: melhoria do tempo de compilação de consultas em até 2,5 vezes com a remoção de partições de metastores do Glue.
- Por padrão, quando UDFs os integrados são passados pelo Hive para o Hive Metastore Server, somente um subconjunto desses incorporados é UDFs passado para o Glue Metastore, já que o Glue suporta apenas operadores de expressão limitados. Se você definir `hive.glue.partition.pruning.client=true`, toda a remoção de partições ocorrerá no lado do cliente. Se você definir `hive.glue.partition.pruning.server=true`, toda a remoção de partições ocorrerá no lado do servidor.

Problemas conhecidos

- As consultas no Hue não funcionam no Amazon EMR 6.4.0 porque o servidor HttpFS do Apache Hadoop está desabilitado por padrão. Para usar o Hue no Amazon EMR 6.4.0, inicie manualmente o servidor HttpFS no nó primário do Amazon EMR usando `sudo systemctl start hadoop-httpfs`, ou use uma etapa do Amazon EMR.
- O atributo Cadernos do Amazon EMR usado com a personificação de usuários do Livy não funciona porque o HttpFS está desabilitado por padrão. Nesse caso, o caderno do EMR não pode se conectar ao cluster que tem a personificação do Livy habilitada. A solução alternativa é iniciar o servidor HttpFS antes de conectar o caderno do EMR ao cluster usando `sudo systemctl start hadoop-httpfs`.
- Na versão 6.4.0 do Amazon EMR, o Phoenix não é compatível com o componente de conectores do Phoenix.
- Para usar as ações do Spark com o Apache Oozie, você deve adicionar a seguinte configuração ao seu arquivo `workflow.xml` do Oozie. Caso contrário, várias bibliotecas críticas, como Hadoop e EMRFS, estarão ausentes do classpath dos executores do Spark que o Oozie inicia.

```
<spark-opts>--conf spark.yarn.populateHadoopClasspath=true</spark-opts>
```

- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.

- O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere / (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.

Versões de componente 6.4.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.1	SDK do Amazon SageMaker Spark
emr-ddb	4.16.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.2.0	Bibliotecas convenientes para o ecossistema do Hadoop.

Componente	Versão	Descrição
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.3.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.18.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	2.1.0	EMR S3Select Connector
emrfs	2.47.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.13.1	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.13.1	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	3.2.1-amzn-4	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	3.2.1-amzn-4	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.2.1-amzn-4	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.2.1-amzn-4	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	3.2.1-amzn-4	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.2.1-amzn-4	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.2.1-amzn-4	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	3.2.1-amzn-4	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.

Componente	Versão	Descrição
hadoop-yarn-nodemanager	3.2.1-amzn-4	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	3.2.1-amzn-4	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hadoop-yarn-timeline-server	3.2.1-amzn-4	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	2.4.4-amzn-0	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	2.4.4-amzn-0	Serviço para atender uma ou mais HBase regiões.
hbase-client	2.4.4-amzn-0	HBase cliente de linha de comando.
hbase-rest-server	2.4.4-amzn-0	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.4.4-amzn-0	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	3.1.2-amzn-5	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.

Componente	Versão	Descrição
hcatalog-server	3.1.2-amzn-5	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	3.1.2-amzn-5	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.2-amzn-5	O cliente da linha de comando do Hive.
hive-hbase	3.1.2-amzn-5	Cliente do Hive-hbase.
hive-metastore-server	3.1.2-amzn-5	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.2-amzn-5	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.8.0-amzn-0	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.8.0-amzn-0	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-trino	0.8.0-amzn-0	Biblioteca de pacotes para execução do Presto com o Hudi.

Componente	Versão	Descrição
hudi-spark	0.8.0-amzn-0	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.9.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	1.4.1	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.1-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mxnet	1.8.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.
nvidia-cuda	10.1.243	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.5.0	Biblioteca de Visão Computaci onal de Código Aberto.

Componente	Versão	Descrição
phoenix-library	5.1.2	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	5.1.2	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.254.1-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.254.1-amzn-0	O serviço que executa partes de uma consulta.
presto-client	0.254.1-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
trino-coordinator	359	Serviço que aceita consultas e gerencia a execução de consultas entre trino-workers.
trino-worker	359	O serviço que executa partes de uma consulta.

Componente	Versão	Descrição
trino-client	359	Cliente de linha de comando do Trino instalado em mestres de espera de um cluster de alta disponibilidade (HA) em que o servidor do Trino não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.0.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	3.1.2-amzn-0	Os clientes da linha de comando do Spark.
spark-history-server	3.1.2-amzn-0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.1.2-amzn-0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.1.2-amzn-0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
spark-rapids	0.4.1	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs

Componente	Versão	Descrição
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.4.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.41+	O servidor Apache HTTP.
zeppelin-server	0.9.0	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.5.7	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.5.7	ZooKeeper cliente de linha de comando.

Classificações de configuração 6.4.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as

classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-6.4.0

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode,

Classificações	Descrição	Ações de reconfiguração
		SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.

Classificações	Descrição	Ações de reconfiguração
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.

Classificações	Descrição	Ações de reconfiguração
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classificações	Descrição	Ações de reconfiguração
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.

Classificações	Descrição	Ações de reconfiguração
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.

Classificações	Descrição	Ações de reconfiguração
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j	Alterar as configurações de log4j.properties no Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.

Classificações	Descrição	Ações de reconfiguração
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
trino-log	Alterar os valores no arquivo log.properties do Trino.	Restarts Trino-Server (for Trino)
trino-config	Alterar os valores no arquivo config.properties do Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Alterar os valores no arquivo password-authenticator.properties do Trino.	Restarts Trino-Server (for Trino)
trino-env	Alterar os valores no arquivo trino-env.sh do Trino.	Restarts Trino-Server (for Trino)
trino-node	Alterar os valores no arquivo node.properties do Trino.	Not available.
trino-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Trino.	Not available.
trino-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Trino.	Not available.
trino-connector-hive	Alterar os valores no arquivo hive.properties do Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Alterar os valores no arquivo jmx.properties do Trino.	Not available.

Classificações	Descrição	Ações de reconfiguração
trino-connector-kafka	Alterar os valores no arquivo kafka.properties do Trino.	Not available.
trino-connector-localfile	Alterar os valores no arquivo localfile.properties do Trino.	Not available.
trino-connector-memory	Alterar os valores no arquivo memory.properties do Trino.	Not available.
trino-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Trino.	Not available.
trino-connector-mysql	Alterar os valores no arquivo mysql.properties do Trino.	Not available.
trino-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Trino.	Not available.
trino-connector-raptor	Alterar os valores no arquivo raptor.properties do Trino.	Not available.
trino-connector-redis	Alterar os valores no arquivo redis.properties do Trino.	Not available.
trino-connector-redshift	Alterar os valores no arquivo redshift.properties do Trino.	Not available.
trino-connector-tpch	Alterar os valores no arquivo tpch.properties do Trino.	Not available.
trino-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Trino.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie and HiveServe r2.
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.

Classificações	Descrição	Ações de reconfiguração
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Versão 6.3.1 do Amazon EMR

Versões 6.3.1 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterEnterpriseGateway, JupyterHubLivy, MXNet, Oozie, Phoenix, PigPresto, PrestoSQL, Spark, SqoopTensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-6.3.1	emr-6.3.0	emr-6.2.1	emr-6.2.0
AWS SDK for Java	1.11.977	1.11.977	1.11.880	1.11.880
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-

	emr-6.3.1	emr-6.3.0	emr-6.2.1	emr-6.2.0
Delta	-	-	-	-
Flink	1.12.1	1.12.1	1.11.2	1.11.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.6	2.2.6	2.2.6-amzn-0	2.2.6-amzn-0
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.7.0-amzn-0	0.7.0-amzn-0	0.6.0-amzn-1	0.6.0-amzn-1
Hue	4.9.0	4.9.0	4.8.0	4.8.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.2.2	1.2.2	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.7.0	1.7.0	1.7.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.0	5.2.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,245,1	0,245,1	0,238,3	0,238,3
Spark	3.1.1	3.1.1	3.0.1	3.0.1

	emr-6.3.1	emr-6.3.0	emr-6.2.1	emr-6.2.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.3.1	2.3.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	350	350	343	343
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas da versão 6.3.1

Esta é uma versão para corrigir problemas com o Amazon EMR Scaling quando ele não consegue up/scale reduzir a escala de um cluster com sucesso ou causa falhas no aplicativo.

Alterações, melhorias e problemas resolvidos

- Corrigido um problema em que as solicitações de escalabilidade falhavam em um cluster grande e altamente utilizado quando os daemons do Amazon EMR no cluster estavam executando atividades de verificação de integridade, como a coleta do estado do nó do YARN e o estado do nó do HDFS. Isso estava acontecendo porque os daemons no cluster não conseguiam comunicar os dados do status de integridade de um nó aos componentes internos do Amazon EMR.
- Aprimorados os daemons do EMR no cluster para rastrear corretamente os estados dos nós quando são reutilizados endereços IP para melhorar a confiabilidade durante operações de escalabilidade.
- SPARK-29683. Corrigido um problema em que ocorriam falhas de trabalho durante a redução da escala verticalmente do cluster, pois o Spark presumia que todos os nós disponíveis estavam na lista de negação.
- YARN-9011. Corrigido um problema em que ocorriam falhas de trabalho devido a uma condição de corrida na desativação do YARN quando o cluster tentava aumentar ou reduzir a escala verticalmente.

- Corrigido problema com falhas de etapas ou tarefas durante a escalabilidade do cluster ao ser garantido que os estados dos nós fossem sempre consistentes entre os daemons do Amazon EMR no cluster e o YARN/HDFS.
- Corrigido um problema em que operações de cluster, como redução de escala verticalmente e envio de etapas, falhavam para clusters do Amazon EMR habilitados com a autenticação Kerberos. Isso ocorreu porque o daemon no cluster do Amazon EMR não renovou o tíquete Kerberos, que é necessário para se comunicar com segurança com a execução no nó primário. HDFS/YARN
- O HTTPS agora está habilitado por padrão para repositórios do Amazon Linux. Se você estiver usando uma política de VPCE do Amazon S3 para restringir o acesso a buckets específicos, deverá adicionar o novo ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` do bucket do Amazon Linux à sua política (substitua `$region` pela região em que o endpoint está situado). Para obter mais informações, consulte esse tópico nos fóruns de AWS discussão. Anúncio: o Amazon Linux 2 agora oferece suporte à capacidade de usar HTTPS ao se conectar a repositórios de pacotes .

Problemas conhecidos

- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
  - O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

  Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.

Versões de componente 6.3.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.1	SDK do Amazon SageMaker Spark
emr-ddb	4.16.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.2.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.2.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.18.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.

Componente	Versão	Descrição
emr-s3-select	2.1.0	EMR S3Select Connector
emrfs	2.46.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.12.1	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.12.1	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	3.2.1-amzn-3.1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	3.2.1-amzn-3.1	O serviço de nível de nó do HDFS para armazenamento de blocos.

Componente	Versão	Descrição
hadoop-hdfs-library	3.2.1-amzn-3.1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.2.1-amzn-3.1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	3.2.1-amzn-3.1	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.2.1-amzn-3.1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.2.1-amzn-3.1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-3.1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.2.1-amzn-3.1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	3.2.1-amzn-3.1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	3.2.1-amzn-3.1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.

Componente	Versão	Descrição
hbase-hmaster	2.2.6-amzn-1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	2.2.6-amzn-1	Serviço para atender uma ou mais HBase regiões.
hbase-client	2.2.6-amzn-1	HBase cliente de linha de comando.
hbase-rest-server	2.2.6-amzn-1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.2.6-amzn-1	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	3.1.2-amzn-4	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.2-amzn-4	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	3.1.2-amzn-4	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.2-amzn-4	O cliente da linha de comando do Hive.
hive-hbase	3.1.2-amzn-4	Cliente do Hive-hbase.

Componente	Versão	Descrição
hive-metastore-server	3.1.2-amzn-4	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.2-amzn-4	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.7.0-amzn-0	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.7.0-amzn-0	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-prestosql	0.7.0-amzn-0	Biblioteca de pacotes para executar o PrestoSQL com o Hudi.
hudi-spark	0.7.0-amzn-0	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.9.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	1.2.2	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.0-incubating	Interface REST para interagir com o Apache Spark

Componente	Versão	Descrição
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mxnet	1.7.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.
nvidia-cuda	10.1.243	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.5.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	5.0.0- HBase -2,0	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	5.0.0- HBase -2,0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.245.1-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.

Componente	Versão	Descrição
presto-worker	0.245.1-amzn-0	O serviço que executa partes de uma consulta.
presto-client	0.245.1-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
prestosql-coordinator	350	O serviço que aceita consultas e gerencia a execução de consultas entre os prestosql-workers.
prestosql-worker	350	O serviço que executa partes de uma consulta.
prestosql-client	350	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.0.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	3.1.1-amzn-0.1	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	3.1.1-amzn-0.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.1.1-amzn-0.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.1.1-amzn-0.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
spark-rapids	0.4.1	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.4.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.41+	O servidor Apache HTTP.
zeppelin-server	0.9.0	O notebook baseado na web que permite um data analytics interativo.

Componente	Versão	Descrição
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 6.3.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-6.3.1

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Classificações	Descrição	Ações de reconfiguração
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.

Classificações	Descrição	Ações de reconfiguração
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.

Classificações	Descrição	Ações de reconfiguração
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j	Alterar as configurações de log4j.properties no Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.

Classificações	Descrição	Ações de reconfiguração
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
prestosql-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-node	Alterar os valores no arquivo node.properties do PrestoSQL.	Not available.

Classificações	Descrição	Ações de reconfiguração
prestosql-connector-blackhole	Alterar os valores no arquivo blackhole.properties do PrestoSQL.	Not available.
prestosql-connector-cassandra	Alterar os valores no arquivo cassandra.properties do PrestoSQL.	Not available.
prestosql-connector-hive	Alterar os valores no arquivo hive.properties do PrestoSQL.	Restarts Presto-Server (for PrestoSQL)
prestosql-connector-jmx	Alterar os valores no arquivo jmx.properties do PrestoSQL.	Not available.
prestosql-connector-kafka	Alterar os valores no arquivo kafka.properties do PrestoSQL.	Not available.
prestosql-connector-localfile	Alterar os valores no arquivo localfile.properties do PrestoSQL.	Not available.
prestosql-connector-memory	Alterar os valores no arquivo memory.properties do PrestoSQL.	Not available.
prestosql-connector-mongodb	Alterar os valores no arquivo mongodb.properties do PrestoSQL.	Not available.
prestosql-connector-mysql	Alterar os valores no arquivo mysql.properties do PrestoSQL.	Not available.
prestosql-connector-postgresql	Alterar os valores no arquivo postgresql.properties do PrestoSQL.	Not available.

Classificações	Descrição	Ações de reconfiguração
prestosql-connector-raptor	Alterar os valores no arquivo raptor.properties do PrestoSQL.	Not available.
prestosql-connector-redis	Alterar os valores no arquivo redis.properties do PrestoSQL.	Not available.
prestosql-connector-redshift	Alterar os valores no arquivo de propriedades redshift do PrestoSQL.	Not available.
prestosql-connector-tpch	Alterar os valores no arquivo tpch.properties do PrestoSQL.	Not available.
prestosql-connector-tpcds	Alterar os valores no arquivo tpcds.properties do PrestoSQL.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.

Classificações	Descrição	Ações de reconfiguração
spark	As configurações administr adas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie and HiveServe r2.

Classificações	Descrição	Ações de reconfiguração
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Versão 6.3.0 do Amazon EMR

Versões 6.3.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterEnterpriseGateway, JupyterHubLivy, MXNet, Oozie, Phoenix, PigPresto, PrestoSQL, Spark, SqoopTensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-6.3.0	emr-6.2.1	emr-6.2.0	emr-6.1.1
AWS SDK for Java	1.11.977	1.11.880	1.11.880	1.11.828
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.12.1	1.11.2	1.11.2	1.11.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.6	2.2.6-amzn-0	2.2.6-amzn-0	2.2.5
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2

	emr-6.3.0	emr-6.2.1	emr-6.2.0	emr-6.1.1
Hudi	0.7.0-amzn-0	0.6.0-amzn-1	0.6.0-amzn-1	0.5.2-incubating-amzn-2
Hue	4.9.0	4.8.0	4.8.0	4.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	-
JupyterHub	1.2.2	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.7.0	1.7.0	1.6.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.0	5.2.0	5.2.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,245,1	0,238,3	0,238,3	0.232
Spark	3.1.1	3.0.1	3.0.1	3.0.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.3.1	2.3.1	2.1.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	350	343	343	338
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas da versão 6.3.0

As notas da versão a seguir incluem informações para a versão 6.3.0 do Amazon EMR. As alterações são referentes à versão 6.2.0.

Data da versão inicial: 12 de maio de 2021

Data da última atualização: 9 de agosto de 2021

Aplicações compatíveis

- AWS SDK para Java versão 1.11.977
- CloudWatch Sink versão 2.1.0
- Conector do DynamoDB versão 4.16.0
- EMRFS versão 2.46.0
- Amazon EMR Goodies versão 3.2.0
- Conector Kinesis versão 3.5.0 do Amazon EMR
- Servidor de registros do Amazon EMR versão 2.0.0
- Scripts versão 2.5.0 do Amazon EMR
- Flink versão 1.12.1
- Ganglia versão 3.7.2
- AWS Glue Hive Metastore Client versão 3.2.0
- Hadoop versão 3.2.1-amzn-3
- HBase versão 2.2.6-amzn-1
- HBase-operator-tools 1.0.0
- HCatalog versão 3.1.2-amzn-0
- Hive versão 3.1.2-amzn-4
- Hudi versão 0.7.0-amzn-0
- Hue versão 4.9.0
- Java JDK versão Corretto-8.282.08.1 (compilação 1.8.0_282-b08)
- JupyterHub versão 1.2.0
- Livy versão 0.7.0-incubating
- MXNet versão 1.7.0
- Oozie versão 5.2.1

- Phoenix versão 5.0.0
- Pig versão 0.17.0
- Presto versão 0.245.1-amzn-0
- PrestoSQL versão 350
- KMS do Apache Ranger (criptografia transparente multi-mestre) versão 2.0.0
- ranger-plugins 2.0.1-amzn-0
- ranger-s3-plugin 1.1.0
- SageMaker SDK do Spark versão 1.4.1
- Scala versão 2.12.10 (VM de servidor OpenJDK de 64 bits, Java 1.8.0_282)
- Spark versão 3.1.1-amzn-0
- spark-rapids 0.4.1
- Sqoop versão 1.4.7
- TensorFlow versão 2.4.1
- Tez versão 0.9.2
- Zeppelin versão 0.9.0
- Zookeeper versão 3.4.14
- Conectores e drivers: DynamoDB Connector 4.16.0

Novos recursos

- O Amazon EMR é compatível com Pontos de Acesso Amazon S3, atributo do amazon S3 que permite que você gerencie facilmente o acesso a data lakes compartilhados. Usando o alias do Ponto de Acesso Amazon S3, você pode simplificar o acesso aos dados em grande escala no Amazon EMR. Você pode usar os pontos de acesso do Amazon S3 com todas as versões do Amazon EMR sem custo adicional em todas as regiões onde o AWS Amazon EMR está disponível. Para saber mais sobre Pontos de Acesso Amazon S3 e aliases de ponto de acesso, consulte Usar um alias em estilo de bucket para seu ponto de acesso no Guia do usuário do Amazon S3.
- Novos parâmetros `DescribeReleaseLabel` e `ListReleaseLabel` da API fornecem detalhes do rótulo de lançamento do Amazon EMR. Você pode listar programaticamente as versões disponíveis na região em que a solicitação de API é executada e listar as aplicações disponíveis para um rótulo de lançamento específico do Amazon EMR. Os parâmetros do rótulo de lançamento também listam as versões do Amazon EMR compatíveis com uma aplicação

específica, como o Spark. Essas informações podem ser usadas para iniciar programaticamente clusters do Amazon EMR. Por exemplo, você pode iniciar um cluster usando a versão mais recente dos resultados de `ListReleaseLabel`. Para obter mais informações, consulte DescribeReleaseLabele ListReleaseLabelsna Referência da API do Amazon EMR.

- Com a versão 6.3.0 do Amazon EMR, você pode iniciar um cluster que se integre nativamente ao Apache Ranger. O Apache Ranger é uma estrutura de código aberto para habilitar, monitorar e gerenciar uma segurança de dados abrangente em toda a plataforma Hadoop. Para obter mais informações, consulte Apache Ranger. Com a integração nativa, você pode trazer seu próprio Apache Ranger para aplicar um controle de acesso detalhado aos dados no Amazon EMR. Consulte Integrar o Amazon EMR com o Apache Ranger no Guia de gerenciamento do Amazon EMR.
- Políticas gerenciadas com escopo definido: para se alinhar às AWS melhores práticas, o Amazon EMR introduziu políticas gerenciadas padrão com escopo do EMR v2 como substitutas das políticas que serão descontinuadas. Consulte as políticas gerenciadas do Amazon EMR.
- Status de suporte do Instance Metadata Service (IMDS) V2: Para o Amazon EMR 6.2 ou posterior, os componentes do Amazon EMR são usados para todas as chamadas do IMDS. IMDSv2 Para chamadas IMDS no código do aplicativo, você pode usar ambos IMDSv1 e IMDSv2, ou configurar o IMDS para uso somente IMDSv2 para aumentar a segurança. Se você desabilitar IMDSv1 em versões anteriores do Amazon EMR 6.x, isso causará uma falha na inicialização do cluster.

Alterações, melhorias e problemas resolvidos

- Esta é uma versão para corrigir problemas com o Amazon EMR Scaling quando ele não consegue up/scale reduzir a escala de um cluster com sucesso ou causa falhas no aplicativo.
- Corrigido um problema em que as solicitações de escalabilidade falhavam em um cluster grande e altamente utilizado quando os daemons do Amazon EMR no cluster estavam executando atividades de verificação de integridade, como a coleta do estado do nó do YARN e o estado do nó do HDFS. Isso estava acontecendo porque os daemons no cluster não conseguiam comunicar os dados do status de integridade de um nó aos componentes internos do Amazon EMR.
- Aprimorados os daemons do EMR no cluster para rastrear corretamente os estados dos nós quando são reutilizados endereços IP para melhorar a confiabilidade durante operações de escalabilidade.
- SPARK-29683. Corrigido um problema em que ocorriam falhas de trabalho durante a redução da escala verticalmente do cluster, pois o Spark presumia que todos os nós disponíveis estavam na lista de negação.

- YARN-9011. Corrigido um problema em que ocorriam falhas de trabalho devido a uma condição de corrida na desativação do YARN quando o cluster tentava aumentar ou reduzir a escala verticalmente.
- Corrigido problema com falhas de etapas ou tarefas durante a escalabilidade do cluster ao ser garantido que os estados dos nós fossem sempre consistentes entre os daemons do Amazon EMR no cluster e o YARN/HDFS.
- Corrigido um problema em que operações de cluster, como redução de escala verticalmente e envio de etapas, falhavam para clusters do Amazon EMR habilitados com a autenticação Kerberos. Isso ocorreu porque o daemon no cluster do Amazon EMR não renovou o tíquete Kerberos, que é necessário para se comunicar com segurança com a execução no nó primário. HDFS/YARN
- As versões mais recentes do Amazon EMR corrigem o problema com um limite menor de “Máximo de arquivos abertos” em relação às versões mais antigas AL2 no Amazon EMR. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR agora incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”.
- O padrão do modo de explicação da IU do Spark SQL foi alterado de `extended` para `formatted` no Spark 3.1. O Amazon EMR o reverteu para `extended` com a finalidade de incluir informações do plano lógico na IU do Spark SQL. Isso pode ser revertido definindo `spark.sql.ui.explainMode` como `formatted`.
- As confirmações a seguir foram transferidas da principal ramificação do Spark.

  - [SPARK-34752][COMPILAÇÃO] Bump Jetty para 9.4.37 para abordar CVE-2020-27223.

  - [SPARK-34534] Corrige a ordem dos blockIDs ao usar para buscar blocos. FetchShuffleBlocks

  - [SPARK-34681] [SQL] Corrigir o erro da junção hash completa externa embaralhada ao ser criado o lado esquerdo com condições diferentes.

  - [SPARK-34497] [SQL] Corrigir os provedores de conexão JDBC integrados para restaurar as alterações no contexto de segurança da JVM.
- Para melhorar a interoperabilidade com o RAPIDs plug-in Nvidia Spark, foi adicionada uma solução alternativa para solucionar um problema que impede que a remoção dinâmica de partições seja acionada ao usar o Nvidia Spark RAPIDs com a execução adaptativa de consultas desativada, consulte o problema #1378 do RAPIDS e o problema # #1386 do RAPIDS. Para obter detalhes sobre a nova configuração

`spark.sql.optimizer.dynamicPartitionPruning.enforceBroadcastReuse`, consulte o Problema n.º 1386 do RAPIDS.

- O algoritmo padrão do confirmador de saída de arquivo foi alterado do algoritmo v2 para o algoritmo v1 no Spark 3.1 de código aberto. Para obter mais informações, consulte este Amazon EMR otimizando a performance do Spark - remoção dinâmica de partições.
- O Amazon EMR reverteu para o algoritmo v2, o padrão usado em versões anteriores do Amazon EMR 6.x, para evitar a regressão de performance. Para restaurar o comportamento de código aberto do Spark 3.1, defina `spark.hadoop.mapreduce.fileoutputcommitter.algorithm.version` como 1. O Spark de código aberto fez essa alteração porque a confirmação de tarefas no algoritmo do confirmador de saída de arquivo v2 não é atômica, o que pode causar um problema de exatidão dos dados de saída em alguns casos. No entanto, a confirmação da tarefa no algoritmo v1 também não é atômica. Em alguns cenários, a confirmação da tarefa inclui uma exclusão realizada antes de uma renomeação. Isso pode resultar em um problema silencioso de correção de dados.
- Corrigidos os problemas de ajuste de escala gerenciado em versões anteriores do Amazon EMR e feitas melhorias para que os índices de falha de aplicações fossem significativamente reduzidos.
- Instalou o AWS Java SDK Bundle em cada novo cluster. Esse é um único jar contendo todos os serviços SDKs e suas dependências, em vez de jars de componentes individuais. Para obter mais informações, consulte Dependência no pacote Java SDK.

Problemas conhecidos

- Para clusters de sub-rede privados do Amazon EMR 6.3.0 e 6.2.0, você não pode acessar a interface do usuário da Web do Ganglia. Você receberá um erro de “acesso negado (403)”. Outros sites UIs, como Spark, Hue, Zeppelin JupyterHub, Livy e Tez, estão funcionando normalmente. O acesso à interface do usuário da Web do Ganglia em clusters de sub-redes públicas também está funcionando normalmente. Para resolver esse problema, reinicie o serviço httpd no nó primário com `sudo systemctl restart httpd`. Esse problema foi corrigido na versão 6.4.0 do Amazon EMR.
- Quando o AWS Glue Data Catalog está ativado, o uso do Spark para acessar um AWS Glue DB com URI de localização de string nula pode falhar. Isso acontece com versões anteriores do Amazon EMR, mas com o SPARK-31709 (https://issues.apache). org/jira/browse/SPARK-31709) faz com que se aplique a mais casos. Por exemplo, ao criar uma tabela no AWS Glue DB padrão cujo URI de localização é uma string nula, `spark.sql("CREATE TABLE mytest (key string) location '/table_path';")` falha com a mensagem “Não é possível criar um

caminho a partir de uma string vazia". Para contornar isso, defina manualmente um URI de localização dos seus bancos de dados AWS Glue e crie tabelas nesses bancos de dados usando o Spark.

- No Amazon EMR 6.3.0, o PrestoSQL foi atualizado da versão 343 para a versão 350. Há duas alterações relacionadas à segurança provenientes do código aberto que se relaciona a essa alteração de versão. O controle de acesso ao catálogo baseado em arquivos é alterado de `deny` para `allow` quando as regras de propriedades de tabela, esquema ou sessão não são definidas. Além disso, o controle de acesso ao sistema baseado em arquivos é alterado para oferecer suporte a arquivos sem regras de catálogo definidas. Nesse caso, todo o acesso aos catálogos é permitido.

  Para obter mais informações, consulte a Versão 344 (9 de outubro de 2020).

- Observe o diretório de usuários do Hadoop (/) home/hadoop) is readable by everyone. It has Unix 755 (drwxr-xr-x) directory permissions to allow read access by frameworks like Hive. You can put files in /home/hadoop e seus subdiretórios, mas esteja ciente das permissões nesses diretórios para proteger informações confidenciais.

- Limite inferior de "Máximo de arquivos abertos" em versões mais antigas AL2 [corrigido em versões mais recentes]. Versões do Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 e emr-6.2.0 são baseadas em versões mais antigas do Amazon Linux 2 (), que AL2 têm uma configuração de limite inferior para "Máximo de arquivos abertos" quando clusters do Amazon EMR são criados com a AMI padrão. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR incluem uma correção permanente com uma configuração mais alta de "Máximo de arquivos abertos". Versões com o limite inferior de arquivos abertos causam o erro "Muitos arquivos abertos" ao ser enviado um trabalho do Spark. Nas versões afetadas, a AMI padrão do Amazon EMR tem uma configuração de ulimit padrão de 4096 para "Máximo de arquivos abertos", que é inferior ao limite de 65536 arquivos na AMI mais recente do Amazon Linux 2. A configuração inferior de ulimit para "Máximo de arquivos abertos" causa falhas em trabalhos do Spark quando o driver e o executor do Spark tentam abrir mais de 4096 arquivos. Para corrigir o problema, o Amazon EMR tem um script de ação de bootstrap (BA) que ajusta a configuração de ulimit na criação do cluster.

  Se você está usando uma versão mais antiga do Amazon EMR que não tem a correção permanente para esse problema, a solução alternativa a seguir permite que você defina explicitamente o ulimit instance-controller para um máximo de 65536 arquivos.

Defina explicitamente um ulimit na linha de comando

1. Edite `/etc/systemd/system/instance-controller.service` para adicionar os seguintes parâmetros à seção Serviço.

   ```
   LimitNOFILE=65536

   LimitNPROC=65536
   ```

2. Reiniciar InstanceController

   ```
   $ sudo systemctl daemon-reload

   $ sudo systemctl restart instance-controller
   ```

Defina um ulimit usando a ação de bootstrap (BA)

Você também pode usar um script de ação de bootstrap (BA) para configurar o ulimit instance-controller para 65536 arquivos na criação do cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- **Important**

  Os clusters do EMR que executam Amazon Linux ou Amazon Linux 2 Amazon Machine Images (AMIs) usam o comportamento padrão do Amazon Linux e não baixam e instalam automaticamente atualizações importantes e críticas do kernel que exigem uma reinicialização. Esse é o mesmo comportamento de outras EC2 instâncias da Amazon que executam o Amazon Linux AMI padrão. Se novas atualizações de software do Amazon Linux que exigem reinicialização (como atualizações do kernel, NVIDIA e CUDA) forem disponibilizadas após o lançamento de uma versão do Amazon EMR, as instâncias de cluster do Amazon EMR que executam a AMI padrão não baixarão nem instalarão essas atualizações automaticamente. Para obter atualizações do kernel, você pode personalizar sua AMI do Amazon EMR para usar a AMI do Amazon Linux mais recente.

- Para usar as ações do Spark com o Apache Oozie, você deve adicionar a seguinte configuração ao seu arquivo `workflow.xml` do Oozie. Caso contrário, várias bibliotecas críticas, como Hadoop e EMRFS, estarão ausentes do classpath dos executores do Spark que o Oozie inicia.

  ```
  <spark-opts>--conf spark.yarn.populateHadoopClasspath=true</spark-opts>
  ```

- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
  - O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

  Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.

Versões de componente 6.3.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.1	SDK do Amazon SageMaker Spark
emr-ddb	4.16.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.2.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.2.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.18.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.

Componente	Versão	Descrição
emr-s3-select	2.1.0	EMR S3Select Connector
emrfs	2.46.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.12.1	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.12.1	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	3.2.1-amzn-3	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	3.2.1-amzn-3	O serviço de nível de nó do HDFS para armazenamento de blocos.

Componente	Versão	Descrição
hadoop-hdfs-library	3.2.1-amzn-3	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.2.1-amzn-3	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	3.2.1-amzn-3	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.2.1-amzn-3	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.2.1-amzn-3	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-3	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.2.1-amzn-3	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	3.2.1-amzn-3	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	3.2.1-amzn-3	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.

Componente	Versão	Descrição
hbase-hmaster	2.2.6-amzn-1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	2.2.6-amzn-1	Serviço para atender uma ou mais HBase regiões.
hbase-client	2.2.6-amzn-1	HBase cliente de linha de comando.
hbase-rest-server	2.2.6-amzn-1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.2.6-amzn-1	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	3.1.2-amzn-4	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.2-amzn-4	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	3.1.2-amzn-4	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.2-amzn-4	O cliente da linha de comando do Hive.
hive-hbase	3.1.2-amzn-4	Cliente do Hive-hbase.

Componente	Versão	Descrição
hive-metastore-server	3.1.2-amzn-4	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.2-amzn-4	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.7.0-amzn-0	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.7.0-amzn-0	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-prestosql	0.7.0-amzn-0	Biblioteca de pacotes para executar o PrestoSQL com o Hudi.
hudi-spark	0.7.0-amzn-0	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.9.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	1.2.2	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.0-incubating	Interface REST para interagir com o Apache Spark

Componente	Versão	Descrição
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mxnet	1.7.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68+	Servidor de banco de dados MariaDB.
nvidia-cuda	10.1.243	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.5.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	5.0.0- HBase -2,0	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	5.0.0- HBase -2,0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.245.1-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.

Componente	Versão	Descrição
presto-worker	0.245.1-amzn-0	O serviço que executa partes de uma consulta.
presto-client	0.245.1-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
prestosql-coordinator	350	O serviço que aceita consultas e gerencia a execução de consultas entre os prestosql-workers.
prestosql-worker	350	O serviço que executa partes de uma consulta.
prestosql-client	350	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.0.2	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	3.1.1-amzn-0	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	3.1.1-amzn-0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.1.1-amzn-0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.1.1-amzn-0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
spark-rapids	0.4.1	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.4.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.41+	O servidor Apache HTTP.
zeppelin-server	0.9.0	O notebook baseado na web que permite um data analytics interativo.

Componente	Versão	Descrição
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 6.3.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-6.3.0

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Classificações	Descrição	Ações de reconfiguração
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.

Classificações	Descrição	Ações de reconfiguração
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.

Classificações	Descrição	Ações de reconfiguração
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j	Alterar as configurações de log4j.properties no Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.

Classificações	Descrição	Ações de reconfiguração
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
prestosql-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-node	Alterar os valores no arquivo node.properties do PrestoSQL.	Not available.

Classificações	Descrição	Ações de reconfiguração
prestosql-connector-blackhole	Alterar os valores no arquivo blackhole.properties do PrestoSQL.	Not available.
prestosql-connector-cassandra	Alterar os valores no arquivo cassandra.properties do PrestoSQL.	Not available.
prestosql-connector-hive	Alterar os valores no arquivo hive.properties do PrestoSQL.	Restarts Presto-Server (for PrestoSQL)
prestosql-connector-jmx	Alterar os valores no arquivo jmx.properties do PrestoSQL.	Not available.
prestosql-connector-kafka	Alterar os valores no arquivo kafka.properties do PrestoSQL.	Not available.
prestosql-connector-localfile	Alterar os valores no arquivo localfile.properties do PrestoSQL.	Not available.
prestosql-connector-memory	Alterar os valores no arquivo memory.properties do PrestoSQL.	Not available.
prestosql-connector-mongodb	Alterar os valores no arquivo mongodb.properties do PrestoSQL.	Not available.
prestosql-connector-mysql	Alterar os valores no arquivo mysql.properties do PrestoSQL.	Not available.
prestosql-connector-postgresql	Alterar os valores no arquivo postgresql.properties do PrestoSQL.	Not available.

Classificações	Descrição	Ações de reconfiguração
prestosql-connector-raptor	Alterar os valores no arquivo raptor.properties do PrestoSQL.	Not available.
prestosql-connector-redis	Alterar os valores no arquivo redis.properties do PrestoSQL.	Not available.
prestosql-connector-redshift	Alterar os valores no arquivo de propriedades redshift do PrestoSQL.	Not available.
prestosql-connector-tpch	Alterar os valores no arquivo tpch.properties do PrestoSQL.	Not available.
prestosql-connector-tpcds	Alterar os valores no arquivo tpcds.properties do PrestoSQL.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.

Classificações	Descrição	Ações de reconfiguração
spark	As configurações administr adas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie and HiveServe r2.

Classificações	Descrição	Ações de reconfiguração
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Versão 6.2.1 do Amazon EMR

Versões 6.2.1 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterEnterpriseGateway, JupyterHubLivy, MXNet, Oozie, Phoenix, PigPresto, PrestoSQL, Spark, SqoopTensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-6.2.1	emr-6.2.0	emr-6.1.1	emr-6.1.0
AWS SDK for Java	1.11.880	1.11.880	1.11.828	1.11.828
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.2	1.11.2	1.11.0	1.11.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.6-amzn-0	2.2.6-amzn-0	2.2.5	2.2.5
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2

	emr-6.2.1	emr-6.2.0	emr-6.1.1	emr-6.1.0
Hudi	0.6.0-amzn-1	0.6.0-amzn-1	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2
Hue	4.8.0	4.8.0	4.7.1	4.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.7.0	1.6.0	1.6.0
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,238,3	0,238,3	0.232	0.232
Spark	3.0.1	3.0.1	3.0.0	3.0.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.3.1	2.3.1	2.1.0	2.1.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	343	343	338	338
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas da versão 6.2.1

Esta é uma versão para corrigir problemas com o Amazon EMR Scaling quando ele não consegue up/scale reduzir a escala de um cluster com sucesso ou causa falhas no aplicativo.

Alterações, melhorias e problemas resolvidos

- Corrigido um problema em que as solicitações de escalabilidade falhavam em um cluster grande e altamente utilizado quando os daemons do Amazon EMR no cluster estavam executando atividades de verificação de integridade, como a coleta do estado do nó do YARN e o estado do nó do HDFS. Isso estava acontecendo porque os daemons no cluster não conseguiam comunicar os dados do status de integridade de um nó aos componentes internos do Amazon EMR.
- Aprimorados os daemons do EMR no cluster para rastrear corretamente os estados dos nós quando são reutilizados endereços IP para melhorar a confiabilidade durante operações de escalabilidade.
- SPARK-29683. Corrigido um problema em que ocorriam falhas de trabalho durante a redução da escala verticalmente do cluster, pois o Spark presumia que todos os nós disponíveis estavam na lista de negação.
- YARN-9011. Corrigido um problema em que ocorriam falhas de trabalho devido a uma condição de corrida na desativação do YARN quando o cluster tentava aumentar ou reduzir a escala verticalmente.
- Corrigido problema com falhas de etapas ou tarefas durante a escalabilidade do cluster ao ser garantido que os estados dos nós fossem sempre consistentes entre os daemons do Amazon EMR no cluster e o YARN/HDFS.
- Corrigido um problema em que operações de cluster, como redução de escala verticalmente e envio de etapas, falhavam para clusters do Amazon EMR habilitados com a autenticação Kerberos. Isso ocorreu porque o daemon no cluster do Amazon EMR não renovou o tíquete Kerberos, que é necessário para se comunicar com segurança com a execução no nó primário. HDFS/YARN
- As versões mais recentes do Amazon EMR corrigem o problema com um limite menor de “Máximo de arquivos abertos” em relação às versões mais antigas AL2 no Amazon EMR. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR agora incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”.
- O HTTPS agora está habilitado por padrão para repositórios do Amazon Linux. Se você estiver usando uma política de VPCE do Amazon S3 para restringir o acesso a buckets específicos,

deverá adicionar o novo ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` do bucket do Amazon Linux à sua política (substitua `$region` pela região em que o endpoint está situado). Para obter mais informações, consulte esse tópico nos fóruns de AWS discussão. Anúncio: o Amazon Linux 2 agora oferece suporte à capacidade de usar HTTPS ao se conectar a repositórios de pacotes .

Problemas conhecidos

- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
  - O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

  Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.

Versões de componente 6.2.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`*

começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.1	SDK do Amazon SageMaker Spark
emr-ddb	4.16.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.1.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.0.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.16.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	2.0.0	EMR S3Select Connector
emrfs	2.44.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.11.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.

Componente	Versão	Descrição
flink-jobmanager-config	1.11.2	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	3.2.1-amzn-2.1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	3.2.1-amzn-2.1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.2.1-amzn-2.1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.2.1-amzn-2.1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.

Componente	Versão	Descrição
hadoop-hdfs-journalnode	3.2.1-amzn-2.1	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.2.1-amzn-2.1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.2.1-amzn-2.1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-2.1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.2.1-amzn-2.1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	3.2.1-amzn-2.1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	3.2.1-amzn-2.1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	2.2.6-amzn-0	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.

Componente	Versão	Descrição
hbase-region-server	2.2.6-amzn-0	Serviço para atender uma ou mais HBase regiões.
hbase-client	2.2.6-amzn-0	HBase cliente de linha de comando.
hbase-rest-server	2.2.6-amzn-0	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.2.6-amzn-0	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	3.1.2-amzn-3	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.2-amzn-3	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	3.1.2-amzn-3	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.2-amzn-3	O cliente da linha de comando do Hive.
hive-hbase	3.1.2-amzn-3	Cliente do Hive-hbase.
hive-metastore-server	3.1.2-amzn-3	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.

Componente	Versão	Descrição
hive-server2	3.1.2-amzn-3	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.6.0-amzn-1	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.6.0-amzn-1	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-prestosql	0.6.0-amzn-1	Biblioteca de pacotes para executar o PrestoSQL com o Hudi.
hudi-spark	0.6.0-amzn-1	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.8.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	1.1.0	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mxnet	1.7.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.

Componente	Versão	Descrição
mariadb-server	5.5.64+	Servidor de banco de dados MariaDB.
nvidia-cuda	10.1.243	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.0	O cliente da linha de comando do Oozie.
oozie-server	5.2.0	O serviço que aceita solicitações de fluxo de trabalho do Oozie.
opencv	4.4.0	Biblioteca de Visão Computacional de Código Aberto.
phoenix-library	5.0.0- HBase -2,0	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	5.0.0- HBase -2,0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.238.3-amzn-1	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.238.3-amzn-1	O serviço que executa partes de uma consulta.

Componente	Versão	Descrição
presto-client	0.238.3-amzn-1	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
prestosql-coordinator	343	O serviço que aceita consultas e gerencia a execução de consultas entre os prestosql-workers.
prestosql-worker	343	O serviço que executa partes de uma consulta.
prestosql-client	343	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.3	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	3.0.1-amzn-0.1	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	3.0.1-amzn-0.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.0.1-amzn-0.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.0.1-amzn-0.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
spark-rapids	0.2.0	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.3.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.41+	O servidor Apache HTTP.
zeppelin-server	0.9.0-preview1	O notebook baseado na web que permite um data analytics interativo.

Componente	Versão	Descrição
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 6.2.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-6.2.1

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
flink-conf	Alterar as configurações do flink-conf.yaml.	Not available.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Not available.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.	Not available.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Not available.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Classificações	Descrição	Ações de reconfiguração
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS ZKFC.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.

Classificações	Descrição	Ações de reconfiguração
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administr adas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.

Classificações	Descrição	Ações de reconfiguração
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j	Alterar as configurações de log4j.properties no Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.

Classificações	Descrição	Ações de reconfiguração
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
prestosql-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-node	Alterar os valores no arquivo node.properties do PrestoSQL.	Not available.

Classificações	Descrição	Ações de reconfiguração
prestosql-connector-blackhole	Alterar os valores no arquivo blackhole.properties do PrestoSQL.	Not available.
prestosql-connector-cassandra	Alterar os valores no arquivo cassandra.properties do PrestoSQL.	Not available.
prestosql-connector-hive	Alterar os valores no arquivo hive.properties do PrestoSQL.	Restarts Presto-Server (for PrestoSQL)
prestosql-connector-jmx	Alterar os valores no arquivo jmx.properties do PrestoSQL.	Not available.
prestosql-connector-kafka	Alterar os valores no arquivo kafka.properties do PrestoSQL.	Not available.
prestosql-connector-localfile	Alterar os valores no arquivo localfile.properties do PrestoSQL.	Not available.
prestosql-connector-memory	Alterar os valores no arquivo memory.properties do PrestoSQL.	Not available.
prestosql-connector-mongodb	Alterar os valores no arquivo mongodb.properties do PrestoSQL.	Not available.
prestosql-connector-mysql	Alterar os valores no arquivo mysql.properties do PrestoSQL.	Not available.
prestosql-connector-postgresql	Alterar os valores no arquivo postgresql.properties do PrestoSQL.	Not available.

Classificações	Descrição	Ações de reconfiguração
prestosql-connector-raptor	Alterar os valores no arquivo raptor.properties do PrestoSQL.	Not available.
prestosql-connector-redis	Alterar os valores no arquivo redis.properties do PrestoSQL.	Not available.
prestosql-connector-redshift	Alterar os valores no arquivo de propriedades redshift do PrestoSQL.	Not available.
prestosql-connector-tpch	Alterar os valores no arquivo tpch.properties do PrestoSQL.	Not available.
prestosql-connector-tpcds	Alterar os valores no arquivo tpcds.properties do PrestoSQL.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.

Classificações	Descrição	Ações de reconfiguração
spark	As configurações administr adas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie.

Classificações	Descrição	Ações de reconfiguração
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Versão 6.2.0 do Amazon EMR

Versões 6.2.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterEnterpriseGateway, JupyterHubLivy, MXNet, Oozie, Phoenix, PigPresto, PrestoSQL, Spark, SqoopTensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-6.2.0	emr-6.1.1	emr-6.1.0	emr-6.0.1
AWS SDK for Java	1.11.880	1.11.828	1.11.828	1.11.711
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.2	1.11.0	1.11.0	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.6-amzn-0	2.2.5	2.2.5	2.2.3
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.6.0-amzn-1	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2	0.5.0-incubating-amzn-1
Hue	4.8.0	4.7.1	4.7.1	4.4.0

	emr-6.2.0	emr-6.1.1	emr-6.1.0	emr-6.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	-	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.0.0
Livy	0.7.0	0.7.0	0.7.0	0.6.0
MXNet	1.7.0	1.6.0	1.6.0	1.5.1
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.2.0	5.1.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	-
Presto	0,238,3	0.232	0.232	0.230
Spark	3.0.1	3.0.0	3.0.0	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	-
TensorFlow	2.3.1	2.1.0	2.1.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	343	338	338	-
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas da versão 6.2.0

As notas da versão a seguir incluem informações para a versão 6.2.0 do Amazon EMR. As alterações são referentes à versão 6.1.0.

Data da versão inicial: 9 de dezembro de 2020

Data da última atualização: 4 de outubro de 2021

Aplicações compatíveis

- AWS SDK para Java versão 1.11.828
- emr-record-server versão 1.7.0
- Flink versão 1.11.2
- Ganglia versão 3.7.2
- Hadoop versão 3.2.1-amzn-1
- HBase versão 2.2.6-amzn-0
- HBase-operator-tools 1.0.0
- HCatalog versão 3.1.2-amzn-0
- Hive versão 3.1.2-amzn-3
- Hudi versão 0.6.0-amzn-1
- Hue versão 4.8.0
- JupyterHub versão 1.1.0
- Livy versão 0.7.0
- MXNet versão 1.7.0
- Oozie versão 5.2.0
- Phoenix versão 5.0.0
- Pig versão 0.17.0
- Presto versão 0.238.3-amzn-1
- PrestoSQL versão 343
- Spark versão 3.0.1-amzn-0
- spark-rapids 0.2.0
- TensorFlow versão 2.3.1
- Zeppelin versão 0.9.0-preview1

- Zookeeper versão 3.4.14
- Conectores e drivers: DynamoDB Connector 4.16.0

Novos recursos

- HBase: a renomeação foi removida na fase de confirmação e o HFile rastreamento persistente foi adicionado. Consulte HFile Rastreamento persistente no Guia de lançamento do Amazon EMR.
- HBase: Backported Crie uma configuração que força o armazenamento em cache dos blocos na compactação.
- PrestoDB: melhorias na remoção dinâmica de partições. O Join Reorder baseado em regras funciona em dados não particionados.
- Políticas gerenciadas com escopo definido: para se alinhar às AWS melhores práticas, o Amazon EMR introduziu políticas gerenciadas padrão com escopo do EMR v2 como substitutas das políticas que serão descontinuadas. Consulte as políticas gerenciadas do Amazon EMR.
- Status de suporte do Instance Metadata Service (IMDS) V2: Para o Amazon EMR 6.2 ou posterior, os componentes do Amazon EMR são usados para todas as chamadas do IMDS. IMDSv2 Para chamadas IMDS no código do aplicativo, você pode usar ambos IMDSv1 e IMDSv2, ou configurar o IMDS para uso somente IMDSv2 para aumentar a segurança. Se você desabilitar IMDSv1 em versões anteriores do Amazon EMR 6.x, isso causará uma falha na inicialização do cluster.

Alterações, melhorias e problemas resolvidos

- Esta é uma versão para corrigir problemas com o Amazon EMR Scaling quando ele não consegue up/scale reduzir a escala de um cluster com sucesso ou causa falhas no aplicativo.
- Corrigido um problema em que as solicitações de escalabilidade falhavam em um cluster grande e altamente utilizado quando os daemons do Amazon EMR no cluster estavam executando atividades de verificação de integridade, como a coleta do estado do nó do YARN e o estado do nó do HDFS. Isso estava acontecendo porque os daemons no cluster não conseguiam comunicar os dados do status de integridade de um nó aos componentes internos do Amazon EMR.
- Aprimorados os daemons do EMR no cluster para rastrear corretamente os estados dos nós quando são reutilizados endereços IP para melhorar a confiabilidade durante operações de escalabilidade.
- SPARK-29683. Corrigido um problema em que ocorriam falhas de trabalho durante a redução da escala verticalmente do cluster, pois o Spark presumia que todos os nós disponíveis estavam na lista de negação.

- YARN-9011. Corrigido um problema em que ocorriam falhas de trabalho devido a uma condição de corrida na desativação do YARN quando o cluster tentava aumentar ou reduzir a escala verticalmente.
- Corrigido problema com falhas de etapas ou tarefas durante a escalabilidade do cluster ao ser garantido que os estados dos nós fossem sempre consistentes entre os daemons do Amazon EMR no cluster e o YARN/HDFS.
- Corrigido um problema em que operações de cluster, como redução de escala verticalmente e envio de etapas, falhavam para clusters do Amazon EMR habilitados com a autenticação Kerberos. Isso ocorreu porque o daemon no cluster do Amazon EMR não renovou o tíquete Kerberos, que é necessário para se comunicar com segurança com a execução no nó primário. HDFS/YARN
- As versões mais recentes do Amazon EMR corrigem o problema com um limite menor de “Máximo de arquivos abertos” em relação às versões mais antigas AL2 no Amazon EMR. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR agora incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”.
- Spark: melhorias de performance no runtime do Spark.

Problemas conhecidos

- O Amazon EMR 6.2 tem permissões incorretas definidas no diretório/etc/cron.d/libinstance-controller-java file in EMR 6.2.0. Permissions on the file are 645 (-rw-r--r-x), when they should be 644 (-rw-r--r--). As a result, Amazon EMR version 6.2 does not log instance-state logs, and the /emr/instance-logs vazio. Esse problema foi corrigido nas versões 6.3.0 e posteriores do Amazon EMR.

  Para contornar esse problema, execute o script a seguir como uma ação de bootstrap na inicialização do cluster.

  ```
  #!/bin/bash
  sudo chmod 644 /etc/cron.d/libinstance-controller-java
  ```

- Para clusters de sub-rede privados do Amazon EMR 6.2.0 e 6.3.0, você não pode acessar a interface do usuário da Web do Ganglia. Você receberá um erro de “acesso negado (403)”. Outros sites UIs, como Spark, Hue, Zeppelin JupyterHub, Livy e Tez, estão funcionando normalmente. O acesso à interface do usuário da Web do Ganglia em clusters de sub-redes públicas também

está funcionando normalmente. Para resolver esse problema, reinicie o serviço httpd no nó primário com `sudo systemctl restart httpd`. Esse problema foi corrigido na versão 6.4.0 do Amazon EMR.

- Há um problema no Amazon EMR 6.2.0 em que o httpd falha continuamente, fazendo com que o Ganglia fique indisponível. Você recebe a mensagem de erro “cannot connect to the server”. Para corrigir um cluster que já está em execução com esse problema, use SSH no nó primário do cluster e adicione a linha `Listen 80` ao arquivo `httpd.conf` localizado em `/etc/httpd/conf/httpd.conf`. Esse problema foi corrigido na versão 6.3.0 do Amazon EMR.
- O HTTPD apresenta falha nos clusters do EMR 6.2.0 quando é usada uma configuração de segurança. Isso faz com que a interface de usuário da aplicação Web do Ganglia fique indisponível. Para acessar a interface de usuário da aplicação Web do Ganglia, adicione `Listen 80` ao arquivo `/etc/httpd/conf/httpd.conf` no nó primário do cluster. Para obter mais informações sobre como se conectar ao cluster, consulte Connect to the Primary Node Using SSH.

  Os Cadernos do EMR também não conseguem estabelecer uma conexão com clusters do EMR 6.2.0 quando você usa uma configuração de segurança. O caderno não conseguirá listar os kernels e enviar trabalhos do Spark. Em vez disso, recomendamos que você use os Cadernos do EMR com outra versão do Amazon EMR.
- Limite inferior de “Máximo de arquivos abertos” em versões mais antigas AL2 [corrigido em versões mais recentes]. Versões do Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 e emr-6.2.0 são baseadas em versões mais antigas do Amazon Linux 2 (), que AL2 têm uma configuração de limite inferior para “Máximo de arquivos abertos” quando clusters do Amazon EMR são criados com a AMI padrão. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”. Versões com o limite inferior de arquivos abertos causam o erro “Muitos arquivos abertos” ao ser enviado um trabalho do Spark. Nas versões afetadas, a AMI padrão do Amazon EMR tem uma configuração de ulimit padrão de 4096 para “Máximo de arquivos abertos”, que é inferior ao limite de 65536 arquivos na AMI mais recente do Amazon Linux 2. A configuração inferior de ulimit para “Máximo de arquivos abertos” causa falhas em trabalhos do Spark quando o driver e o executor do Spark tentam abrir mais de 4096 arquivos. Para corrigir o problema, o Amazon EMR tem um script de ação de bootstrap (BA) que ajusta a configuração de ulimit na criação do cluster.

  Se você está usando uma versão mais antiga do Amazon EMR que não tem a correção permanente para esse problema, a solução alternativa a seguir permite que você defina explicitamente o ulimit instance-controller para um máximo de 65536 arquivos.

Defina explicitamente um ulimit na linha de comando

1. Edite `/etc/systemd/system/instance-controller.service` para adicionar os seguintes parâmetros à seção Serviço.

   ```
   LimitNOFILE=65536
   ```

   ```
   LimitNPROC=65536
   ```

2. Reiniciar InstanceController

   ```
   $ sudo systemctl daemon-reload
   ```

   ```
   $ sudo systemctl restart instance-controller
   ```

Defina um ulimit usando a ação de bootstrap (BA)

Você também pode usar um script de ação de bootstrap (BA) para configurar o ulimit instance-controller para 65536 arquivos na criação do cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- > **Important**
  >
  > O Amazon EMR 6.1.0 e 6.2.0 incluem um problema de performance que pode afetar criticamente todas as operações de inserção, upsert e exclusão do Hudi. Se você planeja usar o Hudi com o Amazon EMR 6.1.0 ou 6.2.0, entre em AWS contato com o suporte para obter um Hudi RPM corrigido.

- > **Important**
  >
  > Os clusters do EMR que executam Amazon Linux ou Amazon Linux 2 Amazon Machine Images (AMIs) usam o comportamento padrão do Amazon Linux e não baixam e instalam automaticamente atualizações importantes e críticas do kernel que exigem uma reinicialização. Esse é o mesmo comportamento de outras EC2 instâncias da Amazon que executam o Amazon Linux AMI padrão. Se novas atualizações de software do Amazon Linux que exigem reinicialização (como atualizações do kernel, NVIDIA e CUDA) forem disponibilizadas após o lançamento de uma versão do Amazon EMR, as instâncias de cluster do Amazon EMR que executam a AMI padrão não baixarão nem instalarão essas atualizações automaticamente. Para obter atualizações do kernel, você pode personalizar sua AMI do Amazon EMR para usar a AMI do Amazon Linux mais recente.

- Os artefatos do Maven do Amazon EMR 6.2.0 não são publicados. Eles serão publicados com uma versão futura do Amazon EMR.
- O HFile rastreamento persistente usando a tabela do sistema HBase storefile não oferece suporte ao recurso de replicação da HBase região. Para obter mais informações sobre a replicação de HBase regiões, consulte Leituras de alta disponibilidade consistentes com o cronograma.
- Diferenças entre as versões de bucketing do Amazon EMR 6.x e do EMR 5.x Hive

  O EMR 5.x usa o OOS Apache Hive 2, enquanto o EMR 6.x usa o OOS Apache Hive 3. O Hive2 de código aberto usa o Bucketing versão 1, enquanto o Hive3 de código aberto usa o Bucketing versão 2. Essa diferença de versão de bucketing entre o Hive 2 (EMR 5.x) e o Hive 3 (EMR 6.x) significa que o hash de bucketing do Hive funciona de uma forma diferente. Veja o exemplo abaixo.

  A tabela a seguir é um exemplo criado no EMR 6.x e no EMR 5.x, respectivamente.

```
-- Using following LOCATION in EMR 6.x
CREATE TABLE test_bucketing (id INT, desc STRING)
```

```
PARTITIONED BY (day STRING)
CLUSTERED BY(id) INTO 128 BUCKETS
LOCATION 's3://your-own-s3-bucket/emr-6-bucketing/';

-- Using following LOCATION in EMR 5.x
LOCATION 's3://your-own-s3-bucket/emr-5-bucketing/';
```

Inserir os mesmos dados no EMR 6.x e no EMR 5.x.

```
INSERT INTO test_bucketing PARTITION (day='01') VALUES(66, 'some_data');
INSERT INTO test_bucketing PARTITION (day='01') VALUES(200, 'some_data');
```

A verificação da localização do S3 mostra que o nome do arquivo de bucketing é diferente, pois a função de hash é diferente entre o EMR 6.x (Hive 3) e o EMR 5.x (Hive 2).

```
[hadoop@ip-10-0-0-122 ~]$ aws s3 ls s3://your-own-s3-bucket/emr-6-bucketing/day=01/
2020-10-21 20:35:16         13 000025_0
2020-10-21 20:35:22         14 000121_0
[hadoop@ip-10-0-0-122 ~]$ aws s3 ls s3://your-own-s3-bucket/emr-5-bucketing/day=01/
2020-10-21 20:32:07         13 000066_0
2020-10-21 20:32:51         14 000072_0
```

Você também pode ver a diferença de versão executando o comando a seguir na CLI do Hive no EMR 6.x. Observe que ele retorna a versão 2 do bucketing.

```
hive> DESCRIBE FORMATTED test_bucketing;
...
Table Parameters:
    bucketing_version       2
...
```

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

Solução:

- SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
- Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
  - O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

  Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.

Versões de componente 6.2.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.1	SDK do Amazon SageMaker Spark
emr-ddb	4.16.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.1.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.0.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.16.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.

Componente	Versão	Descrição
emr-s3-select	2.0.0	EMR S3Select Connector
emrfs	2.44.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.11.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.11.2	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	3.2.1-amzn-2	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	3.2.1-amzn-2	O serviço de nível de nó do HDFS para armazenamento de blocos.

Componente	Versão	Descrição
hadoop-hdfs-library	3.2.1-amzn-2	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.2.1-amzn-2	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	3.2.1-amzn-2	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.2.1-amzn-2	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.2.1-amzn-2	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-2	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.2.1-amzn-2	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	3.2.1-amzn-2	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	3.2.1-amzn-2	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.

Componente	Versão	Descrição
hbase-hmaster	2.2.6-amzn-0	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	2.2.6-amzn-0	Serviço para atender uma ou mais HBase regiões.
hbase-client	2.2.6-amzn-0	HBase cliente de linha de comando.
hbase-rest-server	2.2.6-amzn-0	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.2.6-amzn-0	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	3.1.2-amzn-3	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.2-amzn-3	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	3.1.2-amzn-3	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.2-amzn-3	O cliente da linha de comando do Hive.
hive-hbase	3.1.2-amzn-3	Cliente do Hive-hbase.

Componente	Versão	Descrição
hive-metastore-server	3.1.2-amzn-3	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.2-amzn-3	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.6.0-amzn-1	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.6.0-amzn-1	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-prestosql	0.6.0-amzn-1	Biblioteca de pacotes para executar o PrestoSQL com o Hudi.
hudi-spark	0.6.0-amzn-1	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.8.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	1.1.0	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.0-incubating	Interface REST para interagir com o Apache Spark

Componente	Versão	Descrição
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mxnet	1.7.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.64+	Servidor de banco de dados MariaDB.
nvidia-cuda	10.1.243	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.0	O cliente da linha de comando do Oozie.
oozie-server	5.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.4.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	5.0.0- HBase -2,0	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	5.0.0- HBase -2,0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.238.3-amzn-1	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.

Componente	Versão	Descrição
presto-worker	0.238.3-amzn-1	O serviço que executa partes de uma consulta.
presto-client	0.238.3-amzn-1	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
prestosql-coordinator	343	O serviço que aceita consultas e gerencia a execução de consultas entre os prestosql-workers.
prestosql-worker	343	O serviço que executa partes de uma consulta.
prestosql-client	343	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.3	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	3.0.1-amzn-0	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	3.0.1-amzn-0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.0.1-amzn-0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.0.1-amzn-0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
spark-rapids	0.2.0	Plugin Nvidia Spark RAPIDS que acelera o Apache Spark com o. GPUs
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.3.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.41+	O servidor Apache HTTP.
zeppelin-server	0.9.0-preview1	O notebook baseado na web que permite um data analytics interativo.

Componente	Versão	Descrição
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 6.2.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-6.2.0

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
flink-conf	Alterar as configurações do flink-conf.yaml.	Not available.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Not available.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.	Not available.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Not available.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Classificações	Descrição	Ações de reconfiguração
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	This classification should not be reconfigured.
hdfs-env	Alterar os valores no ambiente do HDFS.	Restarts Hadoop HDFS ZKFC.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.	Restarts Hive WebHCat server.

Classificações	Descrição	Ações de reconfiguração
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat server.
hive	As configurações administr adas pelo Amazon EMR para o Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Not available.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.

Classificações	Descrição	Ações de reconfiguração
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.

Classificações	Descrição	Ações de reconfiguração
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j	Alterar as configurações de log4j.properties no Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.

Classificações	Descrição	Ações de reconfiguração
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
prestosql-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-node	Alterar os valores no arquivo node.properties do PrestoSQL.	Not available.

Classificações	Descrição	Ações de reconfiguração
prestosql-connector-blackhole	Alterar os valores no arquivo blackhole.properties do PrestoSQL.	Not available.
prestosql-connector-cassandra	Alterar os valores no arquivo cassandra.properties do PrestoSQL.	Not available.
prestosql-connector-hive	Alterar os valores no arquivo hive.properties do PrestoSQL.	Restarts Presto-Server (for PrestoSQL)
prestosql-connector-jmx	Alterar os valores no arquivo jmx.properties do PrestoSQL.	Not available.
prestosql-connector-kafka	Alterar os valores no arquivo kafka.properties do PrestoSQL.	Not available.
prestosql-connector-localfile	Alterar os valores no arquivo localfile.properties do PrestoSQL.	Not available.
prestosql-connector-memory	Alterar os valores no arquivo memory.properties do PrestoSQL.	Not available.
prestosql-connector-mongodb	Alterar os valores no arquivo mongodb.properties do PrestoSQL.	Not available.
prestosql-connector-mysql	Alterar os valores no arquivo mysql.properties do PrestoSQL.	Not available.
prestosql-connector-postgresql	Alterar os valores no arquivo postgresql.properties do PrestoSQL.	Not available.

Classificações	Descrição	Ações de reconfiguração
prestosql-connector-raptor	Alterar os valores no arquivo raptor.properties do PrestoSQL.	Not available.
prestosql-connector-redis	Alterar os valores no arquivo redis.properties do PrestoSQL.	Not available.
prestosql-connector-redshift	Alterar os valores no arquivo de propriedades redshift do PrestoSQL.	Not available.
prestosql-connector-tpch	Alterar os valores no arquivo tpch.properties do PrestoSQL.	Not available.
prestosql-connector-tpcds	Alterar os valores no arquivo tpcds.properties do PrestoSQL.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.

Classificações	Descrição	Ações de reconfiguração
spark	As configurações administr adas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restart Oozie.

Classificações	Descrição	Ações de reconfiguração
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.pro perties.	Restarts Zookeeper server.

Versão 6.1.1 do Amazon EMR

Versões 6.1.1 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterHub, LivyMXNet, Oozie, Phoenix, Pig, PrestoPrestoSQL, Spark, Sqoop, TensorFlow, TezZeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
AWS SDK for Java	1.11.828	1.11.828	1.11.711	1.11.711
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.11.0	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.5	2.2.5	2.2.3	2.2.3
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2	0.5.0-incubating-amzn-1	0.5.0-incubating-amzn-1
Hue	4.7.1	4.7.1	4.4.0	4.4.0

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.0.0	1.0.0
Livy	0.7.0	0.7.0	0.6.0	0.6.0
MXNet	1.6.0	1.6.0	1.5.1	1.5.1
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.1.0	5.1.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	-	-
Presto	0.232	0.232	0.230	0.230
Spark	3.0.0	3.0.0	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	-	-
TensorFlow	2.1.0	2.1.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	338	338	-	-
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas da versão 6.1.1

Esta é uma versão para corrigir problemas com o Amazon EMR Scaling quando ele não consegue up/scale reduzir a escala de um cluster com sucesso ou causa falhas no aplicativo.

Alterações, melhorias e problemas resolvidos

- Corrigido um problema em que as solicitações de escalabilidade falhavam em um cluster grande e altamente utilizado quando os daemons do Amazon EMR no cluster estavam executando atividades de verificação de integridade, como a coleta do estado do nó do YARN e o estado do nó do HDFS. Isso estava acontecendo porque os daemons no cluster não conseguiam comunicar os dados do status de integridade de um nó aos componentes internos do Amazon EMR.
- Aprimorados os daemons do EMR no cluster para rastrear corretamente os estados dos nós quando são reutilizados endereços IP para melhorar a confiabilidade durante operações de escalabilidade.
- SPARK-29683. Corrigido um problema em que ocorriam falhas de trabalho durante a redução da escala verticalmente do cluster, pois o Spark presumia que todos os nós disponíveis estavam na lista de negação.
- YARN-9011. Corrigido um problema em que ocorriam falhas de trabalho devido a uma condição de corrida na desativação do YARN quando o cluster tentava aumentar ou reduzir a escala verticalmente.
- Corrigido problema com falhas de etapas ou tarefas durante a escalabilidade do cluster ao ser garantido que os estados dos nós fossem sempre consistentes entre os daemons do Amazon EMR no cluster e o YARN/HDFS.
- Corrigido um problema em que operações de cluster, como redução de escala verticalmente e envio de etapas, falhavam para clusters do Amazon EMR habilitados com a autenticação Kerberos. Isso ocorreu porque o daemon no cluster do Amazon EMR não renovou o tíquete Kerberos, que é necessário para se comunicar com segurança com a execução no nó primário. HDFS/YARN
- As versões mais recentes do Amazon EMR corrigem o problema com um limite menor de “Máximo de arquivos abertos” em relação às versões mais antigas AL2 no Amazon EMR. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR agora incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”.
- O HTTPS agora está habilitado por padrão para repositórios do Amazon Linux. Se você estiver usando uma política de VPCE do Amazon S3 para restringir o acesso a buckets específicos,

deverá adicionar o novo ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` do bucket do Amazon Linux à sua política (substitua `$region` pela região em que o endpoint está situado). Para obter mais informações, consulte esse tópico nos fóruns de AWS discussão. Anúncio: o Amazon Linux 2 agora oferece suporte à capacidade de usar HTTPS ao se conectar a repositórios de pacotes .

Versões de componente 6.1.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.3.0	SDK do Amazon SageMaker Spark
emr-ddb	4.14.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.1.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.

Componente	Versão	Descrição
emr-s3-dist-cp	2.14.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	2.0.0	EMR S3Select Connector
emrfs	2.42.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.11.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	3.2.1-amzn-1.1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	3.2.1-amzn-1.1	O serviço de nível de nó do HDFS para armazenamento de blocos.

Componente	Versão	Descrição
hadoop-hdfs-library	3.2.1-amzn-1.1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.2.1-amzn-1.1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	3.2.1-amzn-1.1	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.2.1-amzn-1.1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.2.1-amzn-1.1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-1.1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.2.1-amzn-1.1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	3.2.1-amzn-1.1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	3.2.1-amzn-1.1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.

Componente	Versão	Descrição
hbase-hmaster	2.2.5	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	2.2.5	Serviço para atender uma ou mais HBase regiões.
hbase-client	2.2.5	HBase cliente de linha de comando.
hbase-rest-server	2.2.5	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.2.5	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	3.1.2-amzn-2	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.2-amzn-2	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	3.1.2-amzn-2	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.2-amzn-2	O cliente da linha de comando do Hive.
hive-hbase	3.1.2-amzn-2	Cliente do Hive-hbase.

Componente	Versão	Descrição
hive-metastore-server	3.1.2-amzn-2	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.2-amzn-2	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.5.2-incubating-amzn-2	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.5.2-incubating-amzn-2	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-prestosql	0.5.2-incubating-amzn-2	Biblioteca de pacotes para executar o PrestoSQL com o Hudi.
hudi-spark	0.5.2-incubating-amzn-2	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.7.1	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	1.1.0	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.0-incubating	Interface REST para interagir com o Apache Spark

Componente	Versão	Descrição
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mxnet	1.6.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.64+	Servidor de banco de dados MariaDB.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.0	O cliente da linha de comando do Oozie.
oozie-server	5.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.3.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	5.0.0- HBase -2,0	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	5.0.0- HBase -2,0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.232	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.

Componente	Versão	Descrição
presto-worker	0.232	O serviço que executa partes de uma consulta.
presto-client	0.232	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
prestosql-coordinator	338	O serviço que aceita consultas e gerencia a execução de consultas entre os prestosql-workers.
prestosql-worker	338	O serviço que executa partes de uma consulta.
prestosql-client	338	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.3	O projeto R para computação estatística
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	3.0.0-amzn-0.1	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	3.0.0-amzn-0.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.0.0-amzn-0.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.0.0-amzn-0.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.1.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.41+	O servidor Apache HTTP.
zeppelin-server	0.9.0-preview1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.

Componente	Versão	Descrição
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 6.1.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-6.1.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.

Classificações	Descrição
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-env	Alterar os valores no ambiente do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.

Classificações	Descrição
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.

Classificações	Descrição
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
hudi-env	Altere os valores no ambiente do Hudi.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.

Classificações	Descrição
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.

Classificações	Descrição
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.

Classificações	Descrição
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
prestosql-log	Alterar os valores no arquivo log.properties do Presto.
prestosql-config	Alterar os valores no arquivo config.properties do Presto.
prestosql-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
prestosql-env	Altere os valores no arquivo presto-env.sh do Presto.
prestosql-node	Alterar os valores no arquivo node.properties do PrestoSQL.
prestosql-connector-blackhole	Alterar os valores no arquivo blackhole.properties do PrestoSQL.

Classificações	Descrição
prestosql-connector-cassandra	Alterar os valores no arquivo cassandra .properties do PrestoSQL.
prestosql-connector-hive	Alterar os valores no arquivo hive.properties do PrestoSQL.
prestosql-connector-jmx	Alterar os valores no arquivo jmx.properties do PrestoSQL.
prestosql-connector-kafka	Alterar os valores no arquivo kafka.properties do PrestoSQL.
prestosql-connector-localfile	Alterar os valores no arquivo localfile.properties do PrestoSQL.
prestosql-connector-memory	Alterar os valores no arquivo memory.pr operties do PrestoSQL.
prestosql-connector-mongodb	Alterar os valores no arquivo mongodb.p roperties do PrestoSQL.
prestosql-connector-mysql	Alterar os valores no arquivo mysql.properties do PrestoSQL.
prestosql-connector-postgresql	Alterar os valores no arquivo postgresql.propert ies do PrestoSQL.
prestosql-connector-raptor	Alterar os valores no arquivo raptor.properties do PrestoSQL.
prestosql-connector-redis	Alterar os valores no arquivo redis.properties do PrestoSQL.
prestosql-connector-redshift	Alterar os valores no arquivo de propriedades redshift do PrestoSQL.
prestosql-connector-tpch	Alterar os valores no arquivo tpch.properties do PrestoSQL.

Classificações	Descrição
prestosql-connector-tpcds	Alterar os valores no arquivo tpcds.properties do PrestoSQL.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.

Classificações	Descrição
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 6.1.0 do Amazon EMR

Versões 6.1.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterHub, LivyMXNet, Oozie, Phoenix, Pig, PrestoPrestoSQL, Spark, Sqoop, TensorFlow, TezZeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
AWS SDK for Java	1.11.828	1.11.828	1.11.711	1.11.711
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.11.0	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.5	2.2.5	2.2.3	2.2.3
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2	0.5.0-incubating-amzn-1	0.5.0-incubating-amzn-1
Hue	4.7.1	4.7.1	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.0.0	1.0.0
Livy	0.7.0	0.7.0	0.6.0	0.6.0

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
MXNet	1.6.0	1.6.0	1.5.1	1.5.1
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.1.0	5.1.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	-	-
Presto	0.232	0.232	0.230	0.230
Spark	3.0.0	3.0.0	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	-	-
TensorFlow	2.1.0	2.1.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	338	338	-	-
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas da versão 6.1.0

As notas da versão a seguir incluem informações para a versão 6.1.0 do Amazon EMR. As alterações são referentes à versão 6.0.0.

Data da versão inicial: 4 de setembro de 2020

Data da última atualização: 15 de outubro de 2020

Aplicações compatíveis

- AWS SDK para Java versão 1.11.828

- Flink versão 1.11.0
- Ganglia versão 3.7.2
- Hadoop versão 3.2.1-amzn-1
- HBase versão 2.2.5
- HBase-operator-tools 1.0.0
- HCatalog versão 3.1.2-amzn-0
- Hive versão 3.1.2-amzn-1
- Hudi versão 0.5.2-incubating
- Hue versão 4.7.1
- JupyterHub versão 1.1.0
- Livy versão 0.7.0
- MXNet versão 1.6.0
- Oozie versão 5.2.0
- Phoenix versão 5.0.0
- Presto versão 0.232
- PrestoSQL versão 338
- Spark versão 3.0.0-amzn-0
- TensorFlow versão 2.1.0
- Zeppelin versão 0.9.0-preview1
- Zookeeper versão 3.4.14
- Conectores e drivers: DynamoDB Connector 4.14.0

Novos recursos

- Os tipos de instância ARM são compatíveis desde o Amazon EMR versão 5.30.0 e do Amazon EMR versão 6.1.0.
- Os tipos de instância de uso geral M6g são compatíveis desde as versões 6.1.0 e 5.30.0 do Amazon EMR. Para obter mais informações, consulte Tipos de instâncias compatíveis no Guia de gerenciamento do Amazon EMR.
- O recurso de grupos de EC2 posicionamento é suportado a partir da versão 5.23.0 do Amazon EMR como uma opção para vários clusters de nós primários. Atualmente, somente os tipos de nós primários são compatíveis com o atributo grupo de posicionamento e a estratégia `SPREAD` é

aplicada a estes nós primários. A estratégia `SPREAD` posiciona um pequeno grupo de instâncias em um hardware subjacente separado para evitar a perda de múltiplos nós primários em caso de falha de hardware. Para obter mais informações, consulte Integração do EMR com o EC2 Placement Group no Guia de gerenciamento do Amazon EMR.

- Ajuste de escala gerenciado: com a versão 6.1.0 do Amazon EMR, é possível habilitar o Ajuste de Escala Gerenciado do Amazon EMR para aumentar ou diminuir automaticamente o número de instâncias ou unidades no cluster com base na workload. O Amazon EMR avalia continuamente as métricas do cluster para tomar decisões de ajuste de escala que otimizam os clusters em termos de custo e velocidade. O ajuste de escala gerenciado também está disponível na versão 5.30.0 e posteriores do Amazon EMR, exceto na versão 6.0.0. Para obter mais informações, consulte Escalar recursos de cluster no Guia de gerenciamento do Amazon EMR.
- O PrestoSQL versão 338 é compatível com o EMR 6.1.0. Para obter mais informações, consulte Presto.
  - O PrestoSQL só é compatível com as versões 6.1.0 e posteriores do EMR e não no EMR 6.0.0 ou no EMR 5.x.
  - O nome da aplicação, `Presto`, continua a ser usado para instalar o PrestoDB em clusters. Para instalar o PrestoSQL em clusters, use o nome da aplicação `PrestoSQL`.
  - Você pode instalar o PrestoDB ou o PrestoSQL, mas não pode instalar os dois em um único cluster. Se o PrestoDB e o PrestoSQL forem especificados na tentativa de criação de um cluster, ocorrerá um erro de validação e a solicitação de criação do cluster falhará.
  - O PrestoSQL é compatível com clusters de mestre único e de vários mestres. Em clusters de vários mestres, é necessário um metastore externo do Hive para executar o PrestoSQL ou o PrestoDB. Consulte Aplicações compatíveis em um cluster do EMR com vários nós primários.
- Suporte à autenticação automática do ECR no Apache Hadoop e no Apache Spark com Docker: os usuários do Spark podem usar imagens do Docker do Docker Hub e do Amazon Elastic Container Registry (Amazon ECR) para definir dependências de ambiente e biblioteca.

  Configure o Docker e execute aplicações do Spark com o Docker usando o Amazon EMR 6.x.
- O EMR é compatível com as transações ACID do Apache Hive: o Amazon EMR 6.1.0 adiciona suporte às transações ACID do Hive para estar em conformidade com as propriedades ACID de um banco de dados. Com esse atributo, você pode executar as operações `INSERT, UPDATE, DELETE,` e `MERGE` em tabelas gerenciadas do Hive com dados no Amazon Simple Storage Service (Amazon S3). Esse é um atributo essencial para casos de uso, como ingestão de streaming, redefinição de dados, atualizações em massa usando MERGE e mudanças lentas de

dimensões. Para obter mais informações, incluindo exemplos de configuração e casos de uso, consulte O Amazon EMR é compatível com as transações ACID do Apache Hive.

Alterações, melhorias e problemas resolvidos

- Esta é uma versão para corrigir problemas com o Amazon EMR Scaling quando ele não consegue up/scale reduzir a escala de um cluster com sucesso ou causa falhas no aplicativo.
- Corrigido um problema em que as solicitações de escalabilidade falhavam em um cluster grande e altamente utilizado quando os daemons do Amazon EMR no cluster estavam executando atividades de verificação de integridade, como a coleta do estado do nó do YARN e o estado do nó do HDFS. Isso estava acontecendo porque os daemons no cluster não conseguiam comunicar os dados do status de integridade de um nó aos componentes internos do Amazon EMR.
- Aprimorados os daemons do EMR no cluster para rastrear corretamente os estados dos nós quando são reutilizados endereços IP para melhorar a confiabilidade durante operações de escalabilidade.
- SPARK-29683. Corrigido um problema em que ocorriam falhas de trabalho durante a redução da escala verticalmente do cluster, pois o Spark presumia que todos os nós disponíveis estavam na lista de negação.
- YARN-9011. Corrigido um problema em que ocorriam falhas de trabalho devido a uma condição de corrida na desativação do YARN quando o cluster tentava aumentar ou reduzir a escala verticalmente.
- Corrigido problema com falhas de etapas ou tarefas durante a escalabilidade do cluster ao ser garantido que os estados dos nós fossem sempre consistentes entre os daemons do Amazon EMR no cluster e o YARN/HDFS.
- Corrigido um problema em que operações de cluster, como redução de escala verticalmente e envio de etapas, falhavam para clusters do Amazon EMR habilitados com a autenticação Kerberos. Isso ocorreu porque o daemon no cluster do Amazon EMR não renovou o tíquete Kerberos, que é necessário para se comunicar com segurança com a execução no nó primário. HDFS/YARN
- As versões mais recentes do Amazon EMR corrigem o problema com um limite menor de “Máximo de arquivos abertos” em relação às versões mais antigas AL2 no Amazon EMR. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR agora incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”.

- O Apache Flink não é compatível com o EMR 6.0.0, mas é compatível com o EMR 6.1.0 com o Flink 1.11.0. Esta é a primeira versão do Flink a oficialmente oferecer suporte ao Hadoop 3. Consulte o Anúncio de versão do Apache Flink 1.11.0.
- O Ganglia foi removido dos pacotes padrão do EMR 6.1.0.

Problemas conhecidos

- Limite inferior de “Máximo de arquivos abertos” em versões mais antigas AL2 [corrigido em versões mais recentes]. Versões do Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 e emr-6.2.0 são baseadas em versões mais antigas do Amazon Linux 2 (), que AL2 têm uma configuração de limite inferior para “Máximo de arquivos abertos” quando clusters do Amazon EMR são criados com a AMI padrão. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”. Versões com o limite inferior de arquivos abertos causam o erro “Muitos arquivos abertos” ao ser enviado um trabalho do Spark. Nas versões afetadas, a AMI padrão do Amazon EMR tem uma configuração de ulimit padrão de 4096 para “Máximo de arquivos abertos”, que é inferior ao limite de 65536 arquivos na AMI mais recente do Amazon Linux 2. A configuração inferior de ulimit para “Máximo de arquivos abertos” causa falhas em trabalhos do Spark quando o driver e o executor do Spark tentam abrir mais de 4096 arquivos. Para corrigir o problema, o Amazon EMR tem um script de ação de bootstrap (BA) que ajusta a configuração de ulimit na criação do cluster.

  Se você está usando uma versão mais antiga do Amazon EMR que não tem a correção permanente para esse problema, a solução alternativa a seguir permite que você defina explicitamente o ulimit instance-controller para um máximo de 65536 arquivos.

  Defina explicitamente um ulimit na linha de comando

  1. Edite `/etc/systemd/system/instance-controller.service` para adicionar os seguintes parâmetros à seção Serviço.

     ```
     LimitNOFILE=65536

     LimitNPROC=65536
     ```

  2. Reiniciar InstanceController

     ```
     $ sudo systemctl daemon-reload
     ```

```
$ sudo systemctl restart instance-controller
```

Defina um ulimit usando a ação de bootstrap (BA)

Você também pode usar um script de ação de bootstrap (BA) para configurar o ulimit instance-controller para 65536 arquivos na criação do cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- **Important**

  O Amazon EMR 6.1.0 e 6.2.0 incluem um problema de performance que pode afetar criticamente todas as operações de inserção, upsert e exclusão do Hudi. Se você planeja usar o Hudi com o Amazon EMR 6.1.0 ou 6.2.0, entre em AWS contato com o suporte para obter um Hudi RPM corrigido.

- Se você definir uma configuração personalizada de coleta de lixo com `spark.driver.extraJavaOptions` e `spark.executor.extraJavaOptions`, isso resultará em falha de driver/executor inicialização com o EMR 6.1 devido à configuração conflitante da coleta de lixo. Em vez disso, com o EMR versão 6.1.0, você deve especificar uma configuração personalizada de coleta de resíduos do Spark para drivers e executores com as propriedades `spark.driver.defaultJavaOptions` e `spark.executor.defaultJavaOptions`. Leia

mais em Ambiente de runtime do Apache Spark e Configurar a coleta de resíduos do Spark no Amazon EMR 6.1.0.

- Usar o Pig com o Oozie (e dentro do Hue, já que o Hue usa ações do Oozie para executar scripts do Pig) gera um erro em que uma biblioteca nativa lzo não pode ser carregada. Essa mensagem de erro é informativa e não impede a execução do Pig.
- Suporte de simultaneidade do Hudi: atualmente, o Hudi não é compatível com gravações simultâneas em uma única tabela do Hudi. Além disso, o Hudi reverte todas as alterações feitas por gravadores em andamento antes de permitir que um novo gravador seja iniciado. As gravações simultâneas podem interferir nesse mecanismo e introduzir condições de corrida, o que pode causar corrupção de dados. Você deve garantir que, como parte do seu fluxo de trabalho de processamento de dados, só exista um gravador do Hudi operando em uma tabela do Hudi em qualquer instante. O Hudi permite vários leitores simultâneos operando na mesma tabela do Hudi.
- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

  Solução:

  - SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
  - Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas

definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

- Há um problema no Amazon EMR 6.1.0 que afeta os clusters que executam o Presto. Depois de um longo período (dias), o cluster pode gerar erros, como “su: failed to execute /bin/bash: Resource temporarily unavailable” ou “shell request failed on channel 0”. Esse problema é causado por um processo interno do Amazon EMR (InstanceController) que está gerando muitos processos leves (LWP), o que acaba fazendo com que o usuário do Hadoop exceda seu limite de nproc. Isso impede que o usuário abra processos adicionais. A solução para esse problema é fazer a atualização para o EMR 6.2.0.

Versões de componente 6.1.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.3.0	SDK do Amazon SageMaker Spark
emr-ddb	4.14.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.1.0	Bibliotecas convenientes para o ecossistema do Hadoop.

Componente	Versão	Descrição
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.14.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	2.0.0	EMR S3Select Connector
emrfs	2.42.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.11.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	3.2.1-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.

Componente	Versão	Descrição
hadoop-hdfs-datanode	3.2.1-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.2.1-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.2.1-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	3.2.1-amzn-1	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.2.1-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.2.1-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	3.2.1-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.2.1-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	3.2.1-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.

Componente	Versão	Descrição
hadoop-yarn-timeline-server	3.2.1-amzn-1	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	2.2.5	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	2.2.5	Serviço para atender a uma ou mais HBase regiões.
hbase-client	2.2.5	HBase cliente de linha de comando.
hbase-rest-server	2.2.5	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.2.5	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	3.1.2-amzn-2	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.2-amzn-2	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	3.1.2-amzn-2	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.

Componente	Versão	Descrição
hive-client	3.1.2-amzn-2	O cliente da linha de comando do Hive.
hive-hbase	3.1.2-amzn-2	Cliente do Hive-hbase.
hive-metastore-server	3.1.2-amzn-2	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.2-amzn-2	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.5.2-incubating-amzn-2	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.5.2-incubating-amzn-2	Biblioteca de pacotes para executar o Presto com o Hudi.
hudi-prestosql	0.5.2-incubating-amzn-2	Biblioteca de pacotes para executar o PrestoSQL com o Hudi.
hudi-spark	0.5.2-incubating-amzn-2	Biblioteca de pacotes para executar o Spark com o Hudi.
hue-server	4.7.1	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	1.1.0	Servidor de vários usuários para blocos de anotações Jupyter

Componente	Versão	Descrição
livy-server	0.7.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mxnet	1.6.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.64+	Servidor de banco de dados MariaDB.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.0	O cliente da linha de comando do Oozie.
oozie-server	5.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.3.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	5.0.0- HBase -2,0	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	5.0.0- HBase -2,0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica

Componente	Versão	Descrição
presto-coordinator	0.232	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.232	O serviço que executa partes de uma consulta.
presto-client	0.232	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
prestosql-coordinator	338	O serviço que aceita consultas e gerencia a execução de consultas entre os prestosql-workers.
prestosql-worker	338	O serviço que executa partes de uma consulta.
prestosql-client	338	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.3	O projeto R para computação estatística

Componente	Versão	Descrição
ranger-kms-server	2.0.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	3.0.0-amzn-0	Os clientes da linha de comando do Spark.
spark-history-server	3.0.0-amzn-0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	3.0.0-amzn-0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	3.0.0-amzn-0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.1.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.41+	O servidor Apache HTTP.
zeppelin-server	0.9.0-preview1	O notebook baseado na web que permite um data analytics interativo.

Componente	Versão	Descrição
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 6.1.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-6.1.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.

Classificações	Descrição
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.

Classificações	Descrição
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-env	Alterar os valores no ambiente do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.

Classificações	Descrição
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.proper ties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.x ml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properti es do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
hudi-env	Altere os valores no ambiente do Hudi.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_n otebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.

Classificações	Descrição
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.

Classificações	Descrição
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.

Classificações	Descrição
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
prestosql-log	Alterar os valores no arquivo log.properties do Presto.
prestosql-config	Alterar os valores no arquivo config.properties do Presto.
prestosql-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.

Classificações	Descrição
prestosql-env	Altere os valores no arquivo presto-env.sh do Presto.
prestosql-node	Alterar os valores no arquivo node.properties do PrestoSQL.
prestosql-connector-blackhole	Alterar os valores no arquivo blackhole.properties do PrestoSQL.
prestosql-connector-cassandra	Alterar os valores no arquivo cassandra.properties do PrestoSQL.
prestosql-connector-hive	Alterar os valores no arquivo hive.properties do PrestoSQL.
prestosql-connector-jmx	Alterar os valores no arquivo jmx.properties do PrestoSQL.
prestosql-connector-kafka	Alterar os valores no arquivo kafka.properties do PrestoSQL.
prestosql-connector-localfile	Alterar os valores no arquivo localfile.properties do PrestoSQL.
prestosql-connector-memory	Alterar os valores no arquivo memory.properties do PrestoSQL.
prestosql-connector-mongodb	Alterar os valores no arquivo mongodb.properties do PrestoSQL.
prestosql-connector-mysql	Alterar os valores no arquivo mysql.properties do PrestoSQL.
prestosql-connector-postgresql	Alterar os valores no arquivo postgresql.properties do PrestoSQL.
prestosql-connector-raptor	Alterar os valores no arquivo raptor.properties do PrestoSQL.

Classificações	Descrição
prestosql-connector-redis	Alterar os valores no arquivo redis.properties do PrestoSQL.
prestosql-connector-redshift	Alterar os valores no arquivo de propriedades redshift do PrestoSQL.
prestosql-connector-tpch	Alterar os valores no arquivo tpch.properties do PrestoSQL.
prestosql-connector-tpcds	Alterar os valores no arquivo tpcds.properties do PrestoSQL.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.

Classificações	Descrição
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 6.0.1 do Amazon EMR

Versões 6.0.1 da aplicação

Essa versão inclui os seguintes aplicativos: GangliaHBase, HCatalog, Hadoop, HiveHudi, Hue, JupyterHub, Livy, MXNet, Oozie, Phoenix, Presto, SparkTensorFlow, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
AWS SDK for Java	1.11.828	1.11.828	1.11.711	1.11.711
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.11.0	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.5	2.2.5	2.2.3	2.2.3
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
Hudi	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2	0.5.0-incubating-amzn-1	0.5.0-incubating-amzn-1
Hue	4.7.1	4.7.1	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.0.0	1.0.0
Livy	0.7.0	0.7.0	0.6.0	0.6.0
MXNet	1.6.0	1.6.0	1.5.1	1.5.1
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.1.0	5.1.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	-	-
Presto	0.232	0.232	0.230	0.230
Spark	3.0.0	3.0.0	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	-	-
TensorFlow	2.1.0	2.1.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	338	338	-	-
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas da versão 6.0.1

Esta é uma versão para corrigir problemas com o Amazon EMR Scaling quando ele não consegue up/scale reduzir a escala de um cluster com sucesso ou causa falhas no aplicativo.

Alterações, melhorias e problemas resolvidos

- Corrigido um problema em que as solicitações de escalabilidade falhavam em um cluster grande e altamente utilizado quando os daemons do Amazon EMR no cluster estavam executando atividades de verificação de integridade, como a coleta do estado do nó do YARN e o estado do nó do HDFS. Isso estava acontecendo porque os daemons no cluster não conseguiam comunicar os dados do status de integridade de um nó aos componentes internos do Amazon EMR.
- Aprimorados os daemons do EMR no cluster para rastrear corretamente os estados dos nós quando são reutilizados endereços IP para melhorar a confiabilidade durante operações de escalabilidade.
- SPARK-29683. Corrigido um problema em que ocorriam falhas de trabalho durante a redução da escala verticalmente do cluster, pois o Spark presumia que todos os nós disponíveis estavam na lista de negação.
- YARN-9011. Corrigido um problema em que ocorriam falhas de trabalho devido a uma condição de corrida na desativação do YARN quando o cluster tentava aumentar ou reduzir a escala verticalmente.
- Corrigido problema com falhas de etapas ou tarefas durante a escalabilidade do cluster ao ser garantido que os estados dos nós fossem sempre consistentes entre os daemons do Amazon EMR no cluster e o YARN/HDFS.
- Corrigido um problema em que operações de cluster, como redução de escala verticalmente e envio de etapas, falhavam para clusters do Amazon EMR habilitados com a autenticação Kerberos. Isso ocorreu porque o daemon no cluster do Amazon EMR não renovou o tíquete Kerberos, que é necessário para se comunicar com segurança com a execução no nó primário. HDFS/YARN
- As versões mais recentes do Amazon EMR corrigem o problema com um limite menor de “Máximo de arquivos abertos” em relação às versões mais antigas AL2 no Amazon EMR. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR agora incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”.
- O HTTPS agora está habilitado por padrão para repositórios do Amazon Linux. Se você estiver usando uma política de VPCE do Amazon S3 para restringir o acesso a buckets específicos,

deverá adicionar o novo ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` do bucket do Amazon Linux à sua política (substitua `$region` pela região em que o endpoint está situado). Para obter mais informações, consulte esse tópico nos fóruns de AWS discussão. Anúncio: o Amazon Linux 2 agora oferece suporte à capacidade de usar HTTPS ao se conectar a repositórios de pacotes .

Versões de componente 6.0.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.2.6	SDK do Amazon SageMaker Spark
emr-ddb	4.14.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	3.0.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.

Componente	Versão	Descrição
emr-s3-dist-cp	2.14.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.5.0	EMR S3Select Connector
emrfs	2.39.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	3.2.1-amzn-0.1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	3.2.1-amzn-0.1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.2.1-amzn-0.1	O cliente de linha de comando e biblioteca do HDFS

Componente	Versão	Descrição
hadoop-hdfs-namenode	3.2.1-amzn-0.1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	3.2.1-amzn-0.1	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.2.1-amzn-0.1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.2.1-amzn-0.1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-0.1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.2.1-amzn-0.1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	3.2.1-amzn-0.1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	3.2.1-amzn-0.1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.

Componente	Versão	Descrição
hbase-hmaster	2.2.3	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	2.2.3	Serviço para atender a uma ou mais HBase regiões.
hbase-client	2.2.3	HBase cliente de linha de comando.
hbase-rest-server	2.2.3	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.2.3	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	3.1.2-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.2-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	3.1.2-amzn-0	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	3.1.2-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	3.1.2-amzn-0	Cliente do Hive-hbase.

Componente	Versão	Descrição
hive-metastore-server	3.1.2-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.2-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.5.0-incubating-amzn-1	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.5.0-incubating-amzn-1	Biblioteca de pacotes para executar o Presto com o Hudi.
hue-server	4.4.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	1.0.0	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.6.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mxnet	1.5.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.

Componente	Versão	Descrição
mariadb-server	5.5.64+	Servidor de banco de dados MariaDB.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.1.0	O cliente da linha de comando do Oozie.
oozie-server	5.1.0	O serviço que aceita solicitações de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computacional de Código Aberto.
phoenix-library	5.0.0- HBase -2,0	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	5.0.0- HBase -2,0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.230	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.230	O serviço que executa partes de uma consulta.

Componente	Versão	Descrição
presto-client	0.230	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
r	3.4.3	O projeto R para computação estatística
spark-client	2.4.4	Os clientes da linha de comando do Spark.
spark-history-server	2.4.4	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.4	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.4.4	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
tensorflow	1.14.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.41+	O servidor Apache HTTP.

Componente	Versão	Descrição
zeppelin-server	0.9.0-SNAPSHOT	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 6.0.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-6.0.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.

Classificações	Descrição
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-env	Alterar os valores no ambiente do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.

Classificações	Descrição
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.

Classificações	Descrição
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.

Classificações	Descrição
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.

Classificações	Descrição
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.

Classificações	Descrição
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.
recordserver-env	Altere os valores no ambiente do EMR. RecordServer
recordserver-conf	Altere os valores no arquivo erver.properties RecordServer do EMR.
recordserver-log4j	Altere os valores no arquivo log4j.properties RecordServer do EMR.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.

Classificações	Descrição
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 6.0.0 do Amazon EMR

Versões 6.0.0 da aplicação

Essa versão inclui os seguintes aplicativos: GangliaHBase, HCatalog, Hadoop, HiveHudi, Hue, JupyterHub, Livy, MXNet, Oozie, Phoenix, Presto, SparkTensorFlow, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
AWS SDK for Java	1.11.828	1.11.828	1.11.711	1.11.711
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.11.0	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.5	2.2.5	2.2.3	2.2.3
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
Hudi	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2	0.5.0-incubating-amzn-1	0.5.0-incubating-amzn-1
Hue	4.7.1	4.7.1	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.0.0	1.0.0
Livy	0.7.0	0.7.0	0.6.0	0.6.0
MXNet	1.6.0	1.6.0	1.5.1	1.5.1
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.1.0	5.1.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	-	-
Presto	0.232	0.232	0.230	0.230
Spark	3.0.0	3.0.0	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	-	-
TensorFlow	2.1.0	2.1.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	338	338	-	-
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas da versão 6.0.0

As notas da versão a seguir incluem informações para a versão 6.0.0 do Amazon EMR.

Data da versão inicial: 10 de março de 2020

Aplicações compatíveis

- AWS SDK para Java versão 1.11.711
- Ganglia versão 3.7.2
- Hadoop versão 3.2.1
- HBase versão 2.2.3
- HCatalog versão 3.1.2
- Hive versão 3.1.2
- Hudi versão 0.5.0 incubadora
- Hue versão 4.4.0
- JupyterHub versão 1.0.0
- Livy versão 0.6.0
- MXNet versão 1.5.1
- Oozie versão 5.1.0
- Phoenix versão 5.0.0
- Presto versão 0.230
- Spark versão 2.4.4
- TensorFlow versão 1.14.0
- Zeppelin versão 0.9.0-SNAPSHOT
- Zookeeper versão 3.4.14
- Conectores e drivers: DynamoDB Connector 4.14.0

> **Note**
>
> Flink, Sqoop, Pig e Mahout não estão disponíveis na versão 6.0.0 do Amazon EMR.

Novos recursos

- Suporte ao runtime do Docker do YARN: aplicações do YARN, como trabalhos do Spark, agora podem ser executados no contexto de um contêiner do Docker. Isso permite que você defina facilmente dependências em uma imagem do Docker sem a necessidade de instalar bibliotecas personalizadas no cluster do Amazon EMR. Para obter mais informações, consulte Configurar a integração do Docker e Executar aplicações do Spark com o Docker usando o Amazon EMR 6.0.0.
- Suporte ao LLAP do Hive – agora o Hive oferece suporte ao modo de execução do LLAP para melhorar o desempenho da consulta. Para obter mais informações, consulte Usar o LLAP do Hive.

Alterações, melhorias e problemas resolvidos

- Esta é uma versão para corrigir problemas com o Amazon EMR Scaling quando ele não consegue up/scale reduzir a escala de um cluster com sucesso ou causa falhas no aplicativo.
- Corrigido um problema em que as solicitações de escalabilidade falhavam em um cluster grande e altamente utilizado quando os daemons do Amazon EMR no cluster estavam executando atividades de verificação de integridade, como a coleta do estado do nó do YARN e o estado do nó do HDFS. Isso estava acontecendo porque os daemons no cluster não conseguiam comunicar os dados do status de integridade de um nó aos componentes internos do Amazon EMR.
- Aprimorados os daemons do EMR no cluster para rastrear corretamente os estados dos nós quando são reutilizados endereços IP para melhorar a confiabilidade durante operações de escalabilidade.
- SPARK-29683. Corrigido um problema em que ocorriam falhas de trabalho durante a redução da escala verticalmente do cluster, pois o Spark presumia que todos os nós disponíveis estavam na lista de negação.
- YARN-9011. Corrigido um problema em que ocorriam falhas de trabalho devido a uma condição de corrida na desativação do YARN quando o cluster tentava aumentar ou reduzir a escala verticalmente.
- Corrigido problema com falhas de etapas ou tarefas durante a escalabilidade do cluster ao ser garantido que os estados dos nós fossem sempre consistentes entre os daemons do Amazon EMR no cluster e o YARN/HDFS.
- Corrigido um problema em que operações de cluster, como redução de escala verticalmente e envio de etapas, falhavam para clusters do Amazon EMR habilitados com a autenticação Kerberos. Isso ocorreu porque o daemon no cluster do Amazon EMR não renovou o tíquete

Kerberos, que é necessário para se comunicar com segurança com a execução no nó primário. HDFS/YARN

- As versões mais recentes do Amazon EMR corrigem o problema com um limite menor de “Máximo de arquivos abertos” em relação às versões mais antigas AL2 no Amazon EMR. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR agora incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”.
- Amazon Linux
  - O Amazon Linux 2 é o sistema operacional da série 6.x do EMR.
  - `systemd` é usado para gerenciamento de serviço em vez de `upstart`, usado no Amazon Linux 1.
- Java Development Kit (JDK)
  - O JDK 8 do Coretto é o JDK padrão da série 6.x do EMR.
- Scala
  - O Scala 2.12 é usado com o Apache Spark e com o Apache Livy.
- Python 3
  - Agora o Python 3 é a versão padrão do Python no EMR.
- Rótulos de nó do YARN
  - A partir do Amazon EMR série 6.x, o recurso de rótulos de nó do YARN é desabilitado por padrão. Os principais processos do aplicativo podem ser executados tanto nos nós core como nos nós de tarefa por padrão. É possível habilitar o recurso de rótulos de nó do YARN configurando as seguintes propriedades: `yarn.node-labels.enabled` e `yarn.node-labels.am.default-node-label-expression`. Para obter mais informações, consulte Noções básicas sobre nós de tarefa, centrais e primários.

Problemas conhecidos

- Limite inferior de “Máximo de arquivos abertos” em versões mais antigas AL2 [corrigido em versões mais recentes]. Versões do Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 e emr-6.2.0 são baseadas em versões mais antigas do Amazon Linux 2 (), que AL2 têm uma configuração de limite inferior para “Máximo de arquivos abertos” quando clusters do Amazon EMR são criados com a AMI padrão. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”. Versões com o limite inferior de

arquivos abertos causam o erro “Muitos arquivos abertos” ao ser enviado um trabalho do Spark. Nas versões afetadas, a AMI padrão do Amazon EMR tem uma configuração de ulimit padrão de 4096 para “Máximo de arquivos abertos”, que é inferior ao limite de 65536 arquivos na AMI mais recente do Amazon Linux 2. A configuração inferior de ulimit para “Máximo de arquivos abertos” causa falhas em trabalhos do Spark quando o driver e o executor do Spark tentam abrir mais de 4096 arquivos. Para corrigir o problema, o Amazon EMR tem um script de ação de bootstrap (BA) que ajusta a configuração de ulimit na criação do cluster.

Se você está usando uma versão mais antiga do Amazon EMR que não tem a correção permanente para esse problema, a solução alternativa a seguir permite que você defina explicitamente o ulimit instance-controller para um máximo de 65536 arquivos.

Defina explicitamente um ulimit na linha de comando

1. Edite `/etc/systemd/system/instance-controller.service` para adicionar os seguintes parâmetros à seção Serviço.

   ```
   LimitNOFILE=65536

   LimitNPROC=65536
   ```

2. Reiniciar InstanceController

   ```
   $ sudo systemctl daemon-reload

   $ sudo systemctl restart instance-controller
   ```

Defina um ulimit usando a ação de bootstrap (BA)

Você também pode usar um script de ação de bootstrap (BA) para configurar o ulimit instance-controller para 65536 arquivos na criação do cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
```

```
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- O shell interativo do Spark PySpark, incluindo o SparkR e o spark-shell, não é compatível com o uso do Docker com bibliotecas adicionais.
- Para usar o Python 3 com a versão 6.0.0 do Amazon EMR, adicione `PATH` a `yarn.nodemanager.env-whitelist`.
- A funcionalidade Live Long and Process (LLAP) não é suportada quando você usa o AWS Glue Data Catalog como metastore do Hive.
- Ao usar o Amazon EMR 6.0.0 com a integração do Spark e do Docker, você precisa configurar as instâncias no cluster com o mesmo tipo de instância e a mesma quantidade de volumes do EBS para evitar falhas quando enviar um trabalho do Spark com runtime do Docker.
- No Amazon EMR 6.0.0, no Amazon HBase S3, o modo de armazenamento é afetado pelo problema do HBASE-24286. HBase o master não pode inicializar quando o cluster é criado usando dados existentes do S3.
- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

  Solução:

  - SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
  - Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versões de componente 6.0.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.2.6	SDK do Amazon SageMaker Spark
emr-ddb	4.14.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.

Componente	Versão	Descrição
emr-goodies	3.0.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.14.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.5.0	EMR S3Select Connector
emrfs	2.39.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	3.2.1-amzn-0	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.

Componente	Versão	Descrição
hadoop-hdfs-datanode	3.2.1-amzn-0	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	3.2.1-amzn-0	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	3.2.1-amzn-0	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	3.2.1-amzn-0	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	3.2.1-amzn-0	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	3.2.1-amzn-0	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	3.2.1-amzn-0	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	3.2.1-amzn-0	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	3.2.1-amzn-0	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.

Componente	Versão	Descrição
hadoop-yarn-timeline-server	3.2.1-amzn-0	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	2.2.3	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	2.2.3	Serviço para atender a uma ou mais HBase regiões.
hbase-client	2.2.3	HBase cliente de linha de comando.
hbase-rest-server	2.2.3	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	2.2.3	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	3.1.2-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	3.1.2-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	3.1.2-amzn-0	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.

Componente	Versão	Descrição
hive-client	3.1.2-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	3.1.2-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	3.1.2-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	3.1.2-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.5.0-incubating-amzn-1	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.5.0-incubating-amzn-1	Biblioteca de pacotes para executar o Presto com o Hudi.
hue-server	4.4.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	1.0.0	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.6.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP

Componente	Versão	Descrição
mxnet	1.5.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.64+	Servidor de banco de dados MariaDB.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.1.0	O cliente da linha de comando do Oozie.
oozie-server	5.1.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	5.0.0- HBase -2,0	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	5.0.0- HBase -2,0	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.230	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.230	O serviço que executa partes de uma consulta.

Componente	Versão	Descrição
presto-client	0.230	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
r	3.4.3	O projeto R para computação estatística
spark-client	2.4.4	Os clientes da linha de comando do Spark.
spark-history-server	2.4.4	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.4	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.4.4	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
tensorflow	1.14.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.41+	O servidor Apache HTTP.

Componente	Versão	Descrição
zeppelin-server	0.9.0-SNAPSHOT	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 6.0.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-6.0.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.

Classificações	Descrição
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-env	Alterar os valores no ambiente do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.

Classificações	Descrição
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive	As configurações administradas pelo Amazon EMR para o Apache Hive.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.

Classificações	Descrição
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.

Classificações	Descrição
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.

Classificações	Descrição
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.

Classificações	Descrição
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.
recordserver-env	Altere os valores no ambiente do EMR. RecordServer
recordserver-conf	Altere os valores no arquivo erver.properties RecordServer do EMR.
recordserver-log4j	Altere os valores no arquivo log4j.properties RecordServer do EMR.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.

Classificações	Descrição
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versões de lançamento 5.x do Amazon EMR

Esta seção contém versões de aplicações, notas da versão, versões de componentes e classificações de configuração disponíveis em cada versão de lançamento 5.x do Amazon EMR.

Ao executar um cluster, você pode escolher entre várias versões do Amazon EMR. Isso permite que você teste e use versões de aplicações que atendem aos seus requisitos de compatibilidade.

Você especifica o número da versão com o rótulo da versão. Os rótulos de versão estão no formato `emr-x.x.x`. Por exemplo, `.emr-7.9.0`

Novas versões do Amazon EMR são disponibilizadas em diferentes regiões durante um período de vários anos, começando com a primeira região na data da versão inicial. A versão mais recente pode não estar disponível em sua região durante esse período.

Para obter uma tabela abrangente das versões de aplicações em cada versão 5.x do Amazon EMR, consulte Versões de aplicações nas versões 5.x do Amazon EMR.

Tópicos

- Versões de aplicações nas versões 5.x do Amazon EMR
- Versão 5.36.2 do Amazon EMR
- Versão 5.36.1 do Amazon EMR
- Versão 5.36.0 do Amazon EMR
- Versão 5.35.0 do Amazon EMR
- Versão 5.34.0 do Amazon EMR
- Versão 5.33.1 do Amazon EMR
- Versão 5.33.0 do Amazon EMR
- Versão 5.32.1 do Amazon EMR
- Versão 5.32.0 do Amazon EMR
- Versão 5.31.1 do Amazon EMR
- Versão 5.31.0 do Amazon EMR
- Versão 5.30.2 do Amazon EMR
- Versão 5.30.1 do Amazon EMR
- Versão 5.30.0 do Amazon EMR
- Versão 5.29.0 do Amazon EMR
- Versão 5.28.1 do Amazon EMR
- Versão 5.28.0 do Amazon EMR
- Versão 5.27.1 do Amazon EMR
- Versão 5.27.0 do Amazon EMR
- Versão 5.26.0 do Amazon EMR

- Versão 5.25.0 do Amazon EMR
- Versão 5.24.1 do Amazon EMR
- Versão 5.24.0 do Amazon EMR
- Versão 5.23.1 do Amazon EMR
- Versão 5.23.0 do Amazon EMR
- Versão 5.22.0 do Amazon EMR
- Versão 5.21.2 do Amazon EMR
- Versão 5.21.1 do Amazon EMR
- Versão 5.21.0 do Amazon EMR
- Versão 5.20.1 do Amazon EMR
- Versão 5.20.0 do Amazon EMR
- Versão 5.19.1 do Amazon EMR
- Versão 5.19.0 do Amazon EMR
- Versão 5.18.1 do Amazon EMR
- Versão 5.18.0 do Amazon EMR
- Versão 5.17.2 do Amazon EMR
- Versão 5.17.1 do Amazon EMR
- Versão 5.17.0 do Amazon EMR
- Versão 5.16.1 do Amazon EMR
- Versão 5.16.0 do Amazon EMR
- Versão 5.15.1 do Amazon EMR
- Versão 5.15.0 do Amazon EMR
- Versão 5.14.2 do Amazon EMR
- Versão 5.14.1 do Amazon EMR
- Versão 5.14.0 do Amazon EMR
- Versão 5.13.1 do Amazon EMR
- Versão 5.13.0 do Amazon EMR
- Versão 5.12.3 do Amazon EMR

- Versão 5.12.2 do Amazon EMR
- Versão 5.12.1 do Amazon EMR
- Versão 5.12.0 do Amazon EMR
- Versão 5.11.4 do Amazon EMR
- Versão 5.11.3 do Amazon EMR
- Versão 5.11.2 do Amazon EMR
- Versão 5.11.1 do Amazon EMR
- Versão 5.11.0 do Amazon EMR
- Versão 5.10.1 do Amazon EMR
- Versão 5.10.0 do Amazon EMR
- Versão 5.9.1 do Amazon EMR
- Versão 5.9.0 do Amazon EMR
- Versão 5.8.3 do Amazon EMR
- Versão 5.8.2 do Amazon EMR
- Versão 5.8.1 do Amazon EMR
- Versão 5.8.0 do Amazon EMR
- Versão 5.7.1 do Amazon EMR
- Versão 5.7.0 do Amazon EMR
- Versão 5.6.1 do Amazon EMR
- Versão 5.6.0 do Amazon EMR
- Versão 5.5.4 do Amazon EMR
- Versão 5.5.3 do Amazon EMR
- Versão 5.5.2 do Amazon EMR
- Versão 5.5.1 do Amazon EMR
- Versão 5.5.0 do Amazon EMR
- Versão 5.4.1 do Amazon EMR
- Versão 5.4.0 do Amazon EMR
- Versão 5.3.2 do Amazon EMR
- Versão 5.3.1 do Amazon EMR

- Versão 5.3.0 do Amazon EMR
- Versão 5.2.3 do Amazon EMR
- Versão 5.2.2 do Amazon EMR
- Versão 5.2.1 do Amazon EMR
- Versão 5.2.0 do Amazon EMR
- Versão 5.1.1 do Amazon EMR
- Versão 5.1.0 do Amazon EMR
- Versão 5.0.3 do Amazon EMR
- Versão 5.0.2 do Amazon EMR
- Versão 5.0.1 do Amazon EMR
- Versão 5.0.0 do Amazon EMR

Versões de aplicações nas versões 5.x do Amazon EMR

Para obter uma tabela abrangente que liste as versões das aplicações disponíveis em cada versão 5.x do Amazon EMR, abra as versões de aplicações nas versões 5.x do Amazon EMR no seu navegador.

Versão 5.36.2 do Amazon EMR

Versões 5.36.2 da aplicação

Essa versão inclui os seguintes aplicativos: DeltaFlinkGanglia, HBase, HCatalogHadoop, Hive, Hudi, Hue, IcebergJupyterEnterpriseGateway, JupyterHub, Livy, MXNet, MahoutOozie, Phoenix, Pig, Presto, SparkSqoop, TensorFlow, TezZeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR

- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.36.2	emr-5.36.1	emr-5.36.0	emr-5.35.0
AWS SDK for Java	1.12.206	1.12.206	1.12.206	1.12.159
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.2	1.14.2	1.14.2	1.14.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.9	2.3.9	2.3.9	2.3.9
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.9	2.3.9	2.3.9	2.3.9
Hudi	0.10.1-amzn-1	0.10.1-amzn-1	0.10.1-amzn-1	0.9.0-amzn-2
Hue	4.10.0	4.10.0	4.10.0	4.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.6.0	2.6.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1

	emr-5.36.2	emr-5.36.1	emr-5.36.0	emr-5.35.0
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.8.0	1.8.0	1.8.0	1.8.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,267	0,267	0,267	0,266
Spark	2.4.8	2.4.8	2.4.8	2.4.8
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.10.0	0.10.0	0.10.0	0.10.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas da versão 5.36.2

As notas da versão a seguir incluem informações sobre a versão 5.36.2 do Amazon EMR. As alterações são referentes à versão 5.36.1. Para obter informações sobre o cronograma da versão, consulte o log de alterações.

Alterações, melhorias e problemas resolvidos

- Essa versão melhora a lógica de redução vertical da escala do cluster para que o Amazon EMR não reduza a escala dos nós centrais abaixo da configuração do fator de replicação do HDFS para

o cluster. Essa melhoria atende aos requisitos de redundância de dados e reduz a chance de uma operação de ajuste de escala ser interrompida.

- Esta versão adiciona um mecanismo de nova tentativa ao fluxo de trabalho de escalabilidade de clusters para clusters do EMR que executam o Presto ou o Trino. Essa melhoria reduz o risco de o redimensionamento do cluster ser executado indefinidamente devido a uma única operação de redimensionamento com falha. Ela também aprimora a utilização dos clusters, porque seu cluster aumenta e reduz a escala verticalmente com mais rapidez.
- Corrige um problema em que as operações de redução vertical da escala do cluster podem ficar paralisadas enquanto o Amazon EMR desativa normalmente um nó central e ele se torna não íntegro antes de ser totalmente desativado.
- Melhora a estabilidade de um nó em um cluster de alta disponibilidade com vários nós primários quando o Amazon EMR reinicia um único nó.
- Otimiza o gerenciamento de registros com o Amazon EMR em execução na Amazon. EC2 Como resultado, é possível ver uma pequena redução nos custos de armazenamento dos logs do cluster.
- Melhora o gerenciamento dos arquivos de log de ZooKeeper transações que são mantidos nos nós primários para minimizar os cenários em que os arquivos de log ultrapassam os limites e interrompem as operações do cluster.
- Corrige um bug raro que pode fazer com que um cluster de alta disponibilidade com vários nós primários falhe por não conseguir se comunicar com o ResourceManager Yarn.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Malásia), Ásia-Pacífico (Mumbai), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Malásia), Europa (Londres), Ásia-Pacífico (Melbourne), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (AWS GovCloud AWS GovCloud Oregon), México (Central) , Oeste dos EUA (Norte da Califórnia), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv), Ásia-Pacífico (Taipei), Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			China (Ningxia), Europa (Zurique)
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Hyderabad), Ásia-Pacífico (Jacarta) Arta), Ásia-Pacífico (Malásia), Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Malásia) , Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (EAU), Canadá (Central) , Oeste do Canadá

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Calgary), (Leste dos EUA), AWS GovCloud (Oeste dos EUA), China (AWS GovCloud Pequim), China (Ningxia), Israel (Tel Aviv), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Centro Leste (Emirados Árabes Unidos), Ásia-Pacífico (Melbourne),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Israel (Tel Aviv), Oeste do Canadá (Calgary), Europa (Zurique)
2.0.2024 001.0	4.14.352	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 816.0	4.14.350	21 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 809.0	4.14.349	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 719.0	4.14.348	25 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 503.0	4.14.343	xxxxxx, 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

Versões de componente 5.36.2

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.1	SDK do Amazon SageMaker Spark
emr-ddb	4.16.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.16.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.5.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.21.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.

Componente	Versão	Descrição
emr-s3-select	1.7.0	EMR S3Select Connector
emrfs	2.51.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.14.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.14.2	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.10.1-amzn-4	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.10.1-amzn-4	O serviço de nível de nó do HDFS para armazenamento de blocos.

Componente	Versão	Descrição
hadoop-hdfs-library	2.10.1-amzn-4	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.10.1-amzn-4	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	2.10.1-amzn-4	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	2.10.1-amzn-4	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.10.1-amzn-4	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.10.1-amzn-4	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.10.1-amzn-4	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.10.1-amzn-4	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.10.1-amzn-4	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.

Componente	Versão	Descrição
hbase-hmaster	1.4.13	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.13	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.13	HBase cliente de linha de comando.
hbase-rest-server	1.4.13	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.13	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2,3.9-amzn-2	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2,3.9-amzn-2	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2,3.9-amzn-2	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2,3.9-amzn-2	O cliente da linha de comando do Hive.
hive-hbase	2,3.9-amzn-2	Cliente do Hive-hbase.

Componente	Versão	Descrição
hive-metastore-server	2,3.9-amzn-2	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2,3.9-amzn-2	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.10.1-amzn-1	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-spark	0.10.1-amzn-1	Biblioteca de pacotes para executar o Spark com o Hudi.
hudi-presto	0.10.1-amzn-1	Biblioteca de pacotes para executar o Presto com o Hudi.
hue-server	4.10.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	1.4.1	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.1-incubating	Interface REST para interagir com o Apache Spark
nginx	1.13.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP

Componente	Versão	Descrição
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.8.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68	O servidor de banco de dados MySQL.
nvidia-cuda	11.0.194	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.5.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.3- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.3- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.267-amzn-1	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.

Componente	Versão	Descrição
presto-worker	0.267-amzn-1	O serviço que executa partes de uma consulta.
presto-client	0.267-amzn-1	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.0.2	O projeto R para computação estatística
ranger-kms-server	1.2.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	2.4.8-amzn-2	Os clientes da linha de comando do Spark.
spark-history-server	2.4.8-amzn-2	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.8-amzn-2	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.4.8-amzn-2	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.

Componente	Versão	Descrição
tensorflow	2.4.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.10.0	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.36.2

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-5.36.2

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

Classificações	Descrição	Ações de reconfiguração
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Not available.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.

Classificações	Descrição	Ações de reconfiguração
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.

Classificações	Descrição	Ações de reconfiguração
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	Should not be reconfigured.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

Classificações	Descrição	Ações de reconfiguração
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.pr operties da HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Restarts HiveServer2 and HiveMetastore.

Classificações	Descrição	Ações de reconfiguração
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j	Alterar as configurações de log4j.properties no Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server.
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server.
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
recordserver-env	Altere os valores no ambiente do EMR. RecordServer	Restarts EMR record server.
recordserver-conf	Altere os valores no arquivo erver.properties RecordServer do EMR.	Restarts EMR record server.
recordserver-log4j	Altere os valores no arquivo log4j.properties RecordServer do EMR.	Restarts EMR record server.
spark	As configurações administr adas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.

Classificações	Descrição	Ações de reconfiguração
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restarts Oozie.
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.

Classificações	Descrição	Ações de reconfiguração
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Log de alterações 5.36.2

Log de alterações da versão 5.36.2 e das notas da versão

Data	Event	Descrição
2024-05-28	Implantação concluída	Amazon EMR 5.36.2 totalmente implantado em todas as regiões compatíveis
2024-05-28	Publicação de documentos	Notas da versão 5.36.2 do Amazon EMR publicadas pela primeira vez
2024-05-20	Lançamento inicial	A versão 5.36.2 do Amazon EMR foi implantada pela primeira vez em regiões comerciais limitadas

Versão 5.36.1 do Amazon EMR

Versões 5.36.1 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, Iceberg, JupyterEnterpriseGatewayJupyterHub, Livy, MXNet, Mahout, OoziePhoenix, Pig, Presto, SparkSqoop, TensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.36.1	emr-5.36.0	emr-5.35.0	emr-5.34.0
AWS SDK for Java	1.12.206	1.12.206	1.12.159	1.11.970
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.2	1.14.2	1.14.2	1.13.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13

	emr-5.36.1	emr-5.36.0	emr-5.35.0	emr-5.34.0
HCatalog	2.3.9	2.3.9	2.3.9	2.3.8
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.9	2.3.9	2.3.9	2.3.8
Hudi	0.10.1-amzn-1	0.10.1-amzn-1	0.9.0-amzn-2	0.9.0-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.9.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.6.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.8.0	1.8.0	1.8.0	1.8.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,267	0,267	0,266	0,261
Spark	2.4.8	2.4.8	2.4.8	2.4.8
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-5.36.1	emr-5.36.0	emr-5.35.0	emr-5.34.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.10.0	0.10.0	0.10.0	0.10.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas da versão 5.36.1

As notas da versão a seguir incluem informações para a versão 5.36.1 do Amazon EMR. As alterações são referentes à versão 5.36.0. Para obter informações sobre o cronograma da versão, consulte o log de alterações.

Alterações, melhorias e problemas resolvidos

- A versão 5.36.1 do Amazon EMR adiciona suporte ao arquivamento de logs no Amazon S3 durante a redução da escala do cluster verticalmente. Nas versões 5.x anteriores, só era possível arquivar arquivos de log no Amazon S3 durante o encerramento do cluster. Essa melhoria garante que os arquivos de log gerados no cluster persistam no Amazon S3 mesmo após o encerramento do nó. Para obter mais informações, consulte Configurar registro em log e depuração do cluster.
- A versão 5.36.1 aprimora o daemon de gerenciamento de logs no cluster para monitorar pastas de log adicionais no cluster do EMR. Essa melhoria minimiza os cenários de utilização excessiva do disco.
- A versão 5.36.1 reinicia automaticamente o daemon de gerenciamento de logs no cluster quando ele é interrompido. Essa melhoria reduz o risco de os nós parecerem não íntegros devido à utilização excessiva do disco.
- A versão 5.36.1 corrige um problema no qual os daemons do Amazon EMR no nó primário manteriam metadados obsoletos para as instâncias encerradas no cluster. A manutenção de dados obsoletos pode fazer com que o uso da CPU e da memória no cluster cresça sem limites e acabe por causar falhas no cluster.
- Para clusters que são lançados com vários nós primários, a versão 5.36.1 corrige um problema em que uma falha de EC2 hardware da Amazon em um dos nós primários pode causar a falha de um segundo nó primário e tornar seu cluster instável.

- Para clusters configurados com criptografia em trânsito, o Ajuste de escala gerenciado agora reconhece dados embaralhados do Spark. dados embaralhados do Spark são dados que o Spark redistribui entre partições para executar operações específicas. Durante a redução da escala verticalmente, o escalonamento gerenciado ignora as instâncias com dados embaralhados. Isso evita novas tentativas de trabalho e recálculos, que são onerosos em termos de preço e performance. Para obter mais informações sobre operações de shuffle, consulte o Guia de programação do Spark.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Europa (Londres)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			, Ásia-Pacífico (Melbourn e), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (Oregon) AWS GovCloud , Oeste dos EUA (Norte da Califórnia), Ásia-Pací fico (AWS GovCloud Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv), Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha) , Europa (Estocolm o), Europa (Zurique) , Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Ásia-Pací fico (Jacarta) Ásia-Pacífico (Melbourne), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central), Oeste do Canadá (Calgary) AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud , (Leste dos EUA) AWS GovCloud , (Oeste dos EUA), China (Pequim), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Centro Leste (Emirados Árabes Unidos), Ásia-Pacífico (Melbourne),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Israel (Tel Aviv), Oeste do Canadá (Calgary), Europa (Zurique)
2.0.2024 001.0	4.14.352	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 816.0	4.14.350	21 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 809.0	4.14.349	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 719.0	4.14.348	25 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 223.0	4.14.336	8 de março de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 131.0	4.14.336	14 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 124.0	4.14.336	7 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 109.0	4.14.334	24 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 218.0	4.14.330	2 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 206.0	4.14.330	22 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 116.0	4.14.328	11 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 101.0	4.14.327	16 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 020.1	4.14.326	7 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 012.1	4.14.326	26 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 926.0	4.14.322	19 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 906.0	4.14.322	4 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 822.0	4.14.322	30 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 808.0	4.14.320	24 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 727.0	4.14.320	14 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 719.0	4.14.320	2 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 628.0	4.14.318	12 de julho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 612.0	4.14.314	23 de junho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 404.1	4.14.311	18 de abril de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central)

Versões de componente 5.36.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a

versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *CommunityVersion*-amzn-*EmrVersion*. O *EmrVersion* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.1	SDK do Amazon SageMaker Spark
emr-ddb	4.16.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.16.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.5.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.21.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.7.0	EMR S3Select Connector
emrfs	2.51.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.

Componente	Versão	Descrição
flink-client	1.14.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.14.2	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.10.1-amzn-4	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.10.1-amzn-4	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.10.1-amzn-4	O cliente de linha de comando e biblioteca do HDFS

Componente	Versão	Descrição
hadoop-hdfs-namenode	2.10.1-amzn-4	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	2.10.1-amzn-4	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	2.10.1-amzn-4	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.10.1-amzn-4	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.10.1-amzn-4	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.10.1-amzn-4	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.10.1-amzn-4	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.10.1-amzn-4	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.

Componente	Versão	Descrição
hbase-hmaster	1.4.13	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.13	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.13	HBase cliente de linha de comando.
hbase-rest-server	1.4.13	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.13	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2,3.9-amzn-2	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2,3.9-amzn-2	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2,3.9-amzn-2	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2,3.9-amzn-2	O cliente da linha de comando do Hive.
hive-hbase	2,3.9-amzn-2	Cliente do Hive-hbase.

Componente	Versão	Descrição
hive-metastore-server	2,3.9-amzn-2	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2,3.9-amzn-2	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.10.1-amzn-1	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-spark	0.10.1-amzn-1	Biblioteca de pacotes para executar o Spark com o Hudi.
hudi-presto	0.10.1-amzn-1	Biblioteca de pacotes para executar o Presto com o Hudi.
hue-server	4.10.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	1.4.1	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.1-incubating	Interface REST para interagir com o Apache Spark
nginx	1.13.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP

Componente	Versão	Descrição
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.8.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68	O servidor de banco de dados MySQL.
nvidia-cuda	11.0.194	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.5.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.3- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.3- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.267-amzn-1	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.

Componente	Versão	Descrição
presto-worker	0.267-amzn-1	O serviço que executa partes de uma consulta.
presto-client	0.267-amzn-1	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.0.2	O projeto R para computação estatística
ranger-kms-server	1.2.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	2.4.8-amzn-2	Os clientes da linha de comando do Spark.
spark-history-server	2.4.8-amzn-2	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.8-amzn-2	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.4.8-amzn-2	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.

Componente	Versão	Descrição
tensorflow	2.4.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.10.0	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.36.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-5.36.1

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

Classificações	Descrição	Ações de reconfiguração
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Not available.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.

Classificações	Descrição	Ações de reconfiguração
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.

Classificações	Descrição	Ações de reconfiguração
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	Should not be reconfigured.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

Classificações	Descrição	Ações de reconfiguração
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.pr operties da HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Restarts HiveServer2 and HiveMetastore.

Classificações	Descrição	Ações de reconfiguração
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
hudi-defaults	Alterar os valores no arquivo hudi-defaults.conf do Hudi.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j	Alterar as configurações de log4j.properties no Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server.
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server.
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
recordserver-env	Altere os valores no ambiente do EMR. RecordServer	Restarts EMR record server.
recordserver-conf	Altere os valores no arquivo erver.properties RecordServer do EMR.	Restarts EMR record server.
recordserver-log4j	Altere os valores no arquivo log4j.properties RecordServer do EMR.	Restarts EMR record server.
spark	As configurações administr adas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.

Classificações	Descrição	Ações de reconfiguração
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restarts Oozie.
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.

Classificações	Descrição	Ações de reconfiguração
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Log de alterações 5.36.1

Log de alterações da versão 5.36.1 e das notas da versão

Data	Event	Descrição
2023-07-26	Atualizar	Novos rótulos de versão `2.0.20230612.0` e `2.0.20230628.0` do sistema operacional.
2023-05-25	Implantação concluída	Amazon EMR 5.36.1 totalmente implantado em todas as regiões compatíveis
2023-05-09	Publicação de documentos	Notas da versão 5.36.1 do Amazon EMR publicadas pela primeira vez
2023-05-04	Lançamento inicial	A versão 5.36.1 do Amazon EMR foi implantada pela primeira vez em regiões comerciais limitadas

Versão 5.36.0 do Amazon EMR

Versões 5.36.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, Iceberg, JupyterEnterpriseGatewayJupyterHub, Livy, MXNet, Mahout, OoziePhoenix, Pig, Presto, SparkSqoop, TensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.36.0	emr-5.35.0	emr-5.34.0	emr-5.33.1
AWS SDK for Java	1.12.206	1.12.159	1.11.970	1.11.970
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.2	1.14.2	1.13.1	1.12.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13

	emr-5.36.0	emr-5.35.0	emr-5.34.0	emr-5.33.1
HCatalog	2.3.9	2.3.9	2.3.8	2.3.7
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.9	2.3.9	2.3.8	2.3.7
Hudi	0.10.1-amzn-1	0.9.0-amzn-2	0.9.0-amzn-0	0.7.0-amzn-1
Hue	4.10.0	4.10.0	4.9.0	4.9.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.2.2
Livy	0.7.1	0.7.1	0.7.1	0.7.0
MXNet	1.8.0	1.8.0	1.8.0	1.7.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.1	5.2.1	5.2.1	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,267	0,266	0,261	0,245,1
Spark	2.4.8	2.4.8	2.4.8	2.4.7
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-5.36.0	emr-5.35.0	emr-5.34.0	emr-5.33.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.10.0	0.10.0	0.10.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas da versão 5.36.0

As notas da versão a seguir incluem informações para a versão 5.36.0 do Amazon EMR. As alterações são referentes à versão 5.35.0.

Data da versão inicial: 15 de junho de 2022

Novos atributos

- A versão 5.36.0 do Amazon EMR adiciona suporte à linguagem de definição de dados (DDL) com o Apache Spark em clusters habilitados para Apache Ranger. Isso permite o uso do Apache Ranger para gerenciar o acesso a operações como criar, alterar e descartar bancos de dados e tabelas de um cluster do Amazon EMR.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

> **Note**
>
> Essa versão não recebe mais atualizações automáticas da AMI, pois foi substituída por uma ou mais versões de patch. A versão de patch é indicada pelo número após o segundo ponto decimal (`6.8.1`). Para ver se você está usando a versão de patch mais recente, verifique as versões disponíveis no Guia de lançamento ou verifique o menu suspenso de versões do Amazon EMR quando criar um cluster no console ou use a ação de API `ListReleaseLabels` ou da CLI `list-release-labels`. Para obter atualizações sobre novas versões, assine o feed RSS na página Novidades.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pací fico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda) , Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia)), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Europa (Londres) , Ásia-Pacífico (Melbourn e), Europa (Milão), Ásia-Pacífico (Tóquio), Oeste dos EUA (Oregon), Oeste dos EUA (Norte da Califórnia), Ásia-Pacífico (Cingapura), Ásia-Pací fico (Sydney), Israel (Tel Aviv), Canadá (Central), Europa (Espanha), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres) , Europa (Milão), Europa (Paris), Europa (Estocolm o), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres) , Europa (Milão), Europa (Paris), Europa (Espanha) , Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Ásia-Pacífico (Jacarta) Ásia-Pacífico (Melbourn e), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central), Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres) , Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pací fico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres) , Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pací fico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres) , Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pací fico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres) , Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pací fico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres) , Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pací fico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres) , Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pací fico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres) , Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa (Zurique), Ásia-Pacífico

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Melbourne), Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres) , Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa (Zurique), Ásia-Pacífico

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Melbourne), Israel (Tel Aviv)
2.0.2023 504.1	4.14.313	16 de maio de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres) , Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 418.0	4.14.311	3 de maio de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres) , Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 404.1	4.14.311	18 de abril de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres) , Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)
2.0.2023 404.0	4.14.311	10 de abril de 2023	Leste dos EUA (Norte da Virgínia), Europa (Paris)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 320.0	4.14.309	30 de março de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres) , Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 307.0	4.14.305	15 de março de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres) , Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 207.0	4.14.304	22 de fevereiro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres) , Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 119.1	4.14.301	3 de fevereiro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres) , Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 210.1	4.14.301	22 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres) , Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 103.3	4.14.296	5 de dezembro de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres) , Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 004.0	4.14.294	2 de novembro de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres) , Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 912.1	4.14.291	7 de outubro de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres) , Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)
2.0.2022 719.0	4.14.287	10 de agosto de 2022	Oeste dos EUA (Norte da Califórnia), Europa (Paris), Europa (Estocolm o), Europa (Frankfurt), Ásia-Pacífico (Mumbai), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 426.0	4.14.281	14 de junho de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres) , Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

Alterações, melhorias e problemas resolvidos

- As atualizações do Amazon EMR 5.36.0 agora oferecem suporte a: aws-sdk 1.12.206, Hadoop 2.10.1-amzn-4, Hive 2.3.9-amzn-2, Hudi 0.10.1-amzn-1, Spark 2.4.8-amzn-2, Presto 0.267-amzn-1, conectores do Amazon Glue 1.18.0, EMRFS 2.51.0.

Problemas conhecidos

- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
  - O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

  Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.
- Com as versões 5.36.0 e 6.6.0 a 6.9.0 do Amazon EMR, os componentes do serviço `SecretAgent` e `RecordServer` podem sofrer perda de dados de log devido a uma configuração incorreta do padrão de nome de arquivo nas propriedades do Log4j2. A configuração incorreta faz com que os componentes gerem somente um arquivo de log por dia. Quando a estratégia de rotação ocorre, ela substitui o arquivo existente em vez de gerar um novo arquivo de log, conforme esperado. Como solução alternativa, use uma ação de bootstrap para gerar arquivos de log a cada hora e acrescentar um número inteiro de incremento automático no nome do arquivo para lidar com a rotação.

  Para as versões 6.6.0 a 6.9.0 do Amazon EMR, use a seguinte ação de bootstrap ao iniciar um cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

  Para o Amazon EMR 5.36.0, use a ação de bootstrap a seguir ao iniciar um cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

Versões de componente 5.36.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.1	SDK do Amazon SageMaker Spark
emr-ddb	4.16.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.16.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.5.0	Ambiente Conda para bloco de anotações do emr, que

Componente	Versão	Descrição
		inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.21.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.7.0	EMR S3Select Connector
emrfs	2.51.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.14.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.14.2	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.

Componente	Versão	Descrição
hadoop-client	2.10.1-amzn-4	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.10.1-amzn-4	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.10.1-amzn-4	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.10.1-amzn-4	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	2.10.1-amzn-4	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	2.10.1-amzn-4	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.10.1-amzn-4	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.10.1-amzn-4	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.10.1-amzn-4	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.10.1-amzn-4	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.10.1-amzn-4	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.4.13	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.13	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.13	HBase cliente de linha de comando.
hbase-rest-server	1.4.13	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.13	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2,3.9-amzn-2	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2,3.9-amzn-2	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	2,3.9-amzn-2	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2,3.9-amzn-2	O cliente da linha de comando do Hive.
hive-hbase	2,3.9-amzn-2	Cliente do Hive-hbase.
hive-metastore-server	2,3.9-amzn-2	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2,3.9-amzn-2	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.10.1-amzn-1	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-spark	0.10.1-amzn-1	Biblioteca de pacotes para executar o Spark com o Hudi.
hudi-presto	0.10.1-amzn-1	Biblioteca de pacotes para executar o Presto com o Hudi.
hue-server	4.10.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	1.4.1	Servidor de vários usuários para blocos de anotações Jupyter

Componente	Versão	Descrição
livy-server	0.7.1-incubating	Interface REST para interagir com o Apache Spark
nginx	1.13.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.8.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68	O servidor de banco de dados MySQL.
nvidia-cuda	11.0.194	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.5.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.3- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.3- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica

Componente	Versão	Descrição
presto-coordinator	0.267-amzn-1	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.267-amzn-1	O serviço que executa partes de uma consulta.
presto-client	0.267-amzn-1	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.0.2	O projeto R para computação estatística
ranger-kms-server	1.2.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	2.4.8-amzn-2	Os clientes da linha de comando do Spark.
spark-history-server	2.4.8-amzn-2	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.8-amzn-2	O mecanismo de execução em memória para o YARN.

Componente	Versão	Descrição
spark-yarn-slave	2.4.8-amzn-2	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.4.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.10.0	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.36.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-5.36.0

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.

Classificações	Descrição	Ações de reconfiguração
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Not available.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.

Classificações	Descrição	Ações de reconfiguração
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.

Classificações	Descrição	Ações de reconfiguração
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	Should not be reconfigured.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

Classificações	Descrição	Ações de reconfiguração
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.pr operties da HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Restarts HiveServer2 and HiveMetastore.

Classificações	Descrição	Ações de reconfiguração
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j	Alterar as configurações de log4j.properties no Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server.

Classificações	Descrição	Ações de reconfiguração
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server.
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
recordserver-env	Altere os valores no ambiente do EMR. RecordServer	Restarts EMR record server.
recordserver-conf	Altere os valores no arquivo erver.properties RecordServer do EMR.	Restarts EMR record server.
recordserver-log4j	Altere os valores no arquivo log4j.properties RecordServer do EMR.	Restarts EMR record server.
spark	As configurações administr adas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.	Restarts Spark history server and Spark thrift server.

Classificações	Descrição	Ações de reconfiguração
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restarts Oozie.
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.

Classificações	Descrição	Ações de reconfiguração
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Versão 5.35.0 do Amazon EMR

Versões 5.35.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, Iceberg, JupyterEnterpriseGatewayJupyterHub, Livy, MXNet, Mahout, OoziePhoenix, Pig, Presto, SparkSqoop, TensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.35.0	emr-5.34.0	emr-5.33.1	emr-5.33.0
AWS SDK for Java	1.12.159	1.11.970	1.11.970	1.11.970
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12

	emr-5.35.0	emr-5.34.0	emr-5.33.1	emr-5.33.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.2	1.13.1	1.12.1	1.12.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.9	2.3.8	2.3.7	2.3.7
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.9	2.3.8	2.3.7	2.3.7
Hudi	0.9.0-amzn-2	0.9.0-amzn-0	0.7.0-amzn-1	0.7.0-amzn-1
Hue	4.10.0	4.9.0	4.9.0	4.9.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.2.2	1.2.2
Livy	0.7.1	0.7.1	0.7.0	0.7.0
MXNet	1.8.0	1.8.0	1.7.0	1.7.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.1	5.2.1	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0

	emr-5.35.0	emr-5.34.0	emr-5.33.1	emr-5.33.0
Presto	0,266	0,261	0,245,1	0,245,1
Spark	2.4.8	2.4.8	2.4.7	2.4.7
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.10.0	0.10.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas da versão 5.35.0

Esta é a nota de lançamento da versão 5.35.0 do Amazon EMR.

As notas da versão a seguir incluem informações para a versão 5.35.0 do Amazon EMR. As alterações são referentes à versão 5.34.0.

Data da versão inicial: 30 de março de 2022

Novos atributos

- As aplicações da versão 5.35 do Amazon EMR que usam o Log4j 1.x e o Log4j 2.x são atualizadas para usar o Log4j 1.2.17 (ou superior) e o Log4j 2.17.1 (ou superior), respectivamente, e não exigem o uso de ações de bootstrap para mitigar os problemas de CVE em versões anteriores. Consulte Abordagem para mitigar o CVE-2021-44228.

Alterações, melhorias e problemas resolvidos

Alterações no Flink

Alterar tipo	Descrição
Atualizações	• Atualizar a versão do flink para 1.14.2. • log4j atualizado para 2.17.1.

Alterações no Hadoop

Alterar tipo	Descrição
Backports de código aberto do Hadoop desde o EMR 5.34.0	• YARN-10438: Manipule o ContainerID nulo no cliente # () RMService getContainerReport • YARN-7266: threads do manipulador de eventos do servidor de linha do tempo bloqueados • YARN-10438: O ATS 1.5 falha ao iniciar se os RollingLevelDb arquivos estiverem corrompidos ou ausentes • HADOOP-13500: sincronizar a iteração do objeto de propriedades de configuração • YARN-10651: CapacityScheduler travou com a entrada NPE. AbstractYarnScheduler updateNodeResource() • HDFS-12221: Substitua o xerces em XmlEditsVisitor • HDFS-16410: Análise insegura de Xml em OfflineEditsXmlLoader
Alterações e correções no Hadoop	• O Tomcat usado em KMS e HttpFS foi atualizado para 8.5.75 • Na FileSystemOptimizedCommitter V2, o marcador de sucesso foi escrito no caminho de saída do commitJob definido durante a criação do committer. Como os caminhos

Alterar tipo	Descrição
	de saída do commitJob e do nível da tarefa podem ser diferentes, o caminho foi corrigido para usar o caminho definido nos arquivos de manifesto. Para trabalhos do Hive, isso faz com que o marcador de sucesso seja gravado corretamente ao serem executada s operações, como partição dinâmica ou UNION ALL.

Alterações no Hive

Alterar tipo	Descrição
O Hive foi atualizado para a versão 2.3.9 de código aberto, incluindo estas correções do JIRA	• HIVE-17155: findConfFile () em HiveConf .java tem alguns problemas com o caminho de configuração • HIVE-24797: desabilitar a validação de valores padrão ao quando analisar esquemas Avro • HIVE-21563: Melhore o desempenho do getEmptyTable Table# desativando o Once registerAllFunctions • HIVE-18147: Os testes podem falhar com o java.net. BindException: Endereço já em uso • HIVE-24608: voltar para get_table no cliente HMS para Hive 2.3.x • HIVE-21200: Vetorização - coluna de data lançando java.lang. UnsupportedOperati onException para parquet • HIVE-19228: remover o uso do commons-h ttpclient 3.x
Backports de código aberto do Hive desde o EMR 5.34.0	• HIVE-19990: falha na consulta com intervalo literal na condição de junção

Alterar tipo	Descrição
	• HIVE-25824: atualizar o branch-2.3 para log4j 2.17.0 • TEZ-4062: o agendamento de tentativas especulativas deverá ser anulado quando a tarefa for concluída • TEZ-4108: NullPointerException durante a execução especulativa, condição de corrida • TEZ-3918: a configuração de tez.task.log.level não funciona
Atualizações e correções do Hive	• Atualizar a versão Log4j para 2.17.1 • Atualizar a versão do ORC para 1.4.3 • Corrigido o impasse devido à discussão de penalidades em ShuffleScheduler
Novos recursos	• Adicionado atributo para imprimir consultas do Hive em logs AM. Ela fica desabilitada por padrão. Sinalizador/Conf.: `tez.am.emr.print.hive.query.in.log` . Status (padrão): FALSO.

Alterações no Oozie

Alterar tipo	Descrição
Backports de código aberto do Oozie desde o EMR 5.34.0	• OOZIE-3652: O lançador Oozie deve tentar novamente a listagem de diretórios quando ocorrer NoSuchFileException

Alterações no Pig

Alterar tipo	Descrição
Atualizações	• log4j atualizado para 1.2.17.

Problemas conhecidos

- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
  - O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

  Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.

Versões de componente 5.35.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.1	SDK do Amazon SageMaker Spark
emr-ddb	4.16.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.15.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.5.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.20.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.7.0	EMR S3Select Connector
emrfs	2.49,0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.14.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.14.2	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager

Componente	Versão	Descrição
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.10.1-amzn-3	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.10.1-amzn-3	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.10.1-amzn-3	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.10.1-amzn-3	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	2.10.1-amzn-3	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.

Componente	Versão	Descrição
hadoop-httpfs-server	2.10.1-amzn-3	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.10.1-amzn-3	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.10.1-amzn-3	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.10.1-amzn-3	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.10.1-amzn-3	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.10.1-amzn-3	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.4.13	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.13	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.13	HBase cliente de linha de comando.

Componente	Versão	Descrição
hbase-rest-server	1.4.13	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.13	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.9-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.9-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.3.9-amzn-0	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.9-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.3.9-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.3.9-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.9-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.

Componente	Versão	Descrição
hudi	0.9.0-amzn-2	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-spark	0.9.0-amzn-2	Biblioteca de pacotes para executar o Spark com o Hudi.
hudi-presto	0.9.0-amzn-2	Biblioteca de pacotes para executar o Presto com o Hudi.
hue-server	4.10.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	1.4.1	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.1-incubating	Interface REST para interagir com o Apache Spark
nginx	1.13.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.8.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68	O servidor de banco de dados MySQL.
nvidia-cuda	10.1.243	Drivers NVIDIA e toolkit CUDA

Componente	Versão	Descrição
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.5.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.3- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.3- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.266-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.266-amzn-0	O serviço que executa partes de uma consulta.
presto-client	0.266-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.

Componente	Versão	Descrição
r	4.0.2	O projeto R para computação estatística
ranger-kms-server	1.2.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	2.4.8-amzn-1	Os clientes da linha de comando do Spark.
spark-history-server	2.4.8-amzn-1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.8-amzn-1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.4.8-amzn-1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.4.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.

Componente	Versão	Descrição
zeppelin-server	0.10.0	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.35.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-5.35.0

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.

Classificações	Descrição	Ações de reconfiguração
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Not available.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.

Classificações	Descrição	Ações de reconfiguração
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.

Classificações	Descrição	Ações de reconfiguração
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	Should not be reconfigured.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

Classificações	Descrição	Ações de reconfiguração
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.pr operties da HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Restarts HiveServer2 and HiveMetastore.

Classificações	Descrição	Ações de reconfiguração
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j	Alterar as configurações de log4j.properties no Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server.

Classificações	Descrição	Ações de reconfiguração
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server.
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
recordserver-env	Altere os valores no ambiente do EMR. RecordServer	Restarts EMR record server.
recordserver-conf	Altere os valores no arquivo erver.properties RecordServer do EMR.	Restarts EMR record server.
recordserver-log4j	Altere os valores no arquivo log4j.properties RecordServer do EMR.	Restarts EMR record server.
spark	As configurações administr adas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.	Restarts Spark history server and Spark thrift server.

Classificações	Descrição	Ações de reconfiguração
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restarts Oozie.
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.

Classificações	Descrição	Ações de reconfiguração
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Versão 5.34.0 do Amazon EMR

Versões 5.34.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterEnterpriseGateway, JupyterHubLivy, MXNet, Mahout, Oozie, PhoenixPig, Presto, Spark, SqoopTensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.34.0	emr-5.33.1	emr-5.33.0	emr-5.32.1
AWS SDK for Java	1.11.970	1.11.970	1.11.970	1.11.890
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12

	emr-5.34.0	emr-5.33.1	emr-5.33.0	emr-5.32.1
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.13.1	1.12.1	1.12.1	1.11.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.8	2.3.7	2.3.7	2.3.7
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.8	2.3.7	2.3.7	2.3.7
Hudi	0.9.0-amzn-0	0.7.0-amzn-1	0.7.0-amzn-1	0.6.0-amzn-0
Hue	4.9.0	4.9.0	4.9.0	4.8.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.2.2	1.2.2	1.1.0
Livy	0.7.1	0.7.0	0.7.0	0.7.0
MXNet	1.8.0	1.7.0	1.7.0	1.7.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.1	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0

	emr-5.34.0	emr-5.33.1	emr-5.33.0	emr-5.32.1
Presto	0,261	0,245,1	0,245,1	0,240,1
Spark	2.4.8	2.4.7	2.4.7	2.4.7
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.3.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.10.0	0.9.0	0.9.0	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas da versão 5.34.0

As notas da versão a seguir incluem informações para a versão 5.34.0 do Amazon EMR. As alterações são referentes à versão 5.33.1.

Data da versão inicial: 20 de janeiro de 2022

Data da versão atualizada: 21 de março de 2022

Novos atributos

- [Ajuste de escala gerenciado] Otimização do ajuste de escala gerenciado de dados embaralhados do Spark: para as versões 5.34.0 e posteriores do Amazon EMR e as versões 6.4.0 e posteriores do EMR, o ajuste de escala gerenciado agora reconhece dados embaralhados do Spark (dados que o Spark redistribui entre partições para executar operações específicas). Para obter mais informações sobre operações de shuffle, consulte Usar ajuste de escala gerenciado do EMR no Amazon EMR no Guia de gerenciamento do Amazon EMR e no Guia de programação do Spark.
- [Hudi] Melhorias para simplificar a configuração do Hudi. Desabilitado o controle de simultaneidade otimista por padrão.

Alterações, melhorias e problemas resolvidos

- Esta é uma versão para corrigir problemas com o Amazon EMR Scaling quando ele não consegue up/scale reduzir a escala de um cluster com sucesso ou causa falhas no aplicativo.
- Anteriormente, a reinicialização manual do gerenciador de recursos em um cluster multimestre fazia com que os daemons do Amazon EMR no cluster, como o Zookeeper, recarregassem todos os nós anteriormente desativados ou perdidos no arquivo znode do Zookeeper. Isso fez com que os limites padrão fossem excedidos em determinadas situações. O Amazon EMR agora remove os registros de nós desativados ou perdidos há mais de uma hora do arquivo do Zookeeper e os limites internos foram aumentados.
- Corrigido um problema em que as solicitações de escalabilidade falhavam em um cluster grande e altamente utilizado quando os daemons do Amazon EMR no cluster estavam executando atividades de verificação de integridade, como a coleta do estado do nó do YARN e o estado do nó do HDFS. Isso estava acontecendo porque os daemons no cluster não conseguiam comunicar os dados do status de integridade de um nó aos componentes internos do Amazon EMR.
- Aprimorados os daemons do EMR no cluster para rastrear corretamente os estados dos nós quando são reutilizados endereços IP para melhorar a confiabilidade durante operações de escalabilidade.
- SPARK-29683. Corrigido um problema em que ocorriam falhas de trabalho durante a redução da escala verticalmente do cluster, pois o Spark presumia que todos os nós disponíveis estavam na lista de negação.
- YARN-9011. Corrigido um problema em que ocorriam falhas de trabalho devido a uma condição de corrida na desativação do YARN quando o cluster tentava aumentar ou reduzir a escala verticalmente.
- Corrigido problema com falhas de etapas ou tarefas durante a escalabilidade do cluster ao ser garantido que os estados dos nós fossem sempre consistentes entre os daemons do Amazon EMR no cluster e o YARN/HDFS.
- Corrigido um problema em que operações de cluster, como redução de escala verticalmente e envio de etapas, falhavam para clusters do Amazon EMR habilitados com a autenticação Kerberos. Isso ocorreu porque o daemon no cluster do Amazon EMR não renovou o tíquete Kerberos, que é necessário para se comunicar com segurança com a execução no nó primário. HDFS/YARN
- Zeppelin atualizado para a versão 0.10.0.
- Livy Fix: atualizado para 0.7.1

- Melhoria da performance do Spark: executores heterogêneos são desabilitados quando determinados valores de configuração do Spark são substituídos no EMR 5.34.0.
- WebHDFS e o servidor HttpFS estão desabilitados por padrão. Você pode reabilitar o WebHDFS usando a configuração do Hadoop, `dfs.webhdfs.enabled`. O servidor HttpFS pode ser iniciado usando `sudo systemctl start hadoop-httpfs`.

Problemas conhecidos

- O atributo Cadernos do Amazon EMR usado com a personificação de usuários do Livy não funciona porque o HttpFS está desabilitado por padrão. Nesse caso, o caderno do EMR não pode se conectar ao cluster que tem a personificação do Livy habilitada. A solução alternativa é iniciar o servidor HttpFS antes de conectar o caderno do EMR ao cluster usando `sudo systemctl start hadoop-httpfs`.
- As consultas no Hue não funcionam no Amazon EMR 6.4.0 porque o servidor HttpFS do Apache Hadoop está desabilitado por padrão. Para usar o Hue no Amazon EMR 6.4.0, inicie manualmente o servidor HttpFS no nó primário do Amazon EMR usando `sudo systemctl start hadoop-httpfs`, ou use uma etapa do Amazon EMR.
- O atributo Cadernos do Amazon EMR usado com a personificação de usuários do Livy não funciona porque o HttpFS está desabilitado por padrão. Nesse caso, o caderno do EMR não pode se conectar ao cluster que tem a personificação do Livy habilitada. A solução alternativa é iniciar o servidor HttpFS antes de conectar o caderno do EMR ao cluster usando `sudo systemctl start hadoop-httpfs`.
- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
  - O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.

Versões de componente 5.34.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.1	SDK do Amazon SageMaker Spark
emr-ddb	4.16.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.14.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.4.0	Ambiente Conda para bloco de anotações do emr, que

Componente	Versão	Descrição
		inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.18.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.7.0	EMR S3Select Connector
emrfs	2.48.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.13.1	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.13.1	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.

Componente	Versão	Descrição
hadoop-client	2.10.1-amzn-2	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.10.1-amzn-2	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.10.1-amzn-2	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.10.1-amzn-2	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	2.10.1-amzn-2	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	2.10.1-amzn-2	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.10.1-amzn-2	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.10.1-amzn-2	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.10.1-amzn-2	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.10.1-amzn-2	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.10.1-amzn-2	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.4.13	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.13	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.13	HBase cliente de linha de comando.
hbase-rest-server	1.4.13	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.13	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.8-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.8-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	2.3.8-amzn-0	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.8-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.3.8-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.3.8-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.8-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.9.0-amzn-0	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-spark	0.9.0-amzn-0	Biblioteca de pacotes para executar o Spark com o Hudi.
hudi-presto	0.9.0-amzn-0	Biblioteca de pacotes para executar o Presto com o Hudi.
hue-server	4.9.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	1.4.1	Servidor de vários usuários para blocos de anotações Jupyter

Componente	Versão	Descrição
livy-server	0.7.1-incubating	Interface REST para interagir com o Apache Spark
nginx	1.13.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.8.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68	O servidor de banco de dados MySQL.
nvidia-cuda	10.1.243	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.1	O cliente da linha de comando do Oozie.
oozie-server	5.2.1	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.5.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.3- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.3- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica

Componente	Versão	Descrição
presto-coordinator	0.261-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.261-amzn-0	O serviço que executa partes de uma consulta.
presto-client	0.261-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.0.2	O projeto R para computação estatística
ranger-kms-server	1.2.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	2.4.8-amzn-0	Os clientes da linha de comando do Spark.
spark-history-server	2.4.8-amzn-0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.8-amzn-0	O mecanismo de execução em memória para o YARN.

Componente	Versão	Descrição
spark-yarn-slave	2.4.8-amzn-0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.4.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.10.0	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.34.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-5.34.0

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.

Classificações	Descrição	Ações de reconfiguração
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Not available.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.

Classificações	Descrição	Ações de reconfiguração
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.

Classificações	Descrição	Ações de reconfiguração
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	Should not be reconfigured.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

Classificações	Descrição	Ações de reconfiguração
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.pr operties da HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Restarts HiveServer2 and HiveMetastore.

Classificações	Descrição	Ações de reconfiguração
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j	Alterar as configurações de log4j.properties no Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server.

Classificações	Descrição	Ações de reconfiguração
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server.
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
recordserver-env	Altere os valores no ambiente do EMR. RecordServer	Restarts EMR record server.
recordserver-conf	Altere os valores no arquivo erver.properties RecordServer do EMR.	Restarts EMR record server.
recordserver-log4j	Altere os valores no arquivo log4j.properties RecordServer do EMR.	Restarts EMR record server.
spark	As configurações administr adas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.	Restarts Spark history server and Spark thrift server.

Classificações	Descrição	Ações de reconfiguração
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restarts Oozie.
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.

Classificações	Descrição	Ações de reconfiguração
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Versão 5.33.1 do Amazon EMR

Versões 5.33.1 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterEnterpriseGateway, JupyterHubLivy, MXNet, Mahout, Oozie, PhoenixPig, Presto, Spark, SqoopTensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.33.1	emr-5.33.0	emr-5.32.1	emr-5.32.0
AWS SDK for Java	1.11.970	1.11.970	1.11.890	1.11.890
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12

	emr-5.33.1	emr-5.33.0	emr-5.32.1	emr-5.32.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.12.1	1.12.1	1.11.2	1.11.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.7	2.3.7	2.3.7	2.3.7
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.7	2.3.7	2.3.7	2.3.7
Hudi	0.7.0-amzn-1	0.7.0-amzn-1	0.6.0-amzn-0	0.6.0-amzn-0
Hue	4.9.0	4.9.0	4.8.0	4.8.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.2.2	1.2.2	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.7.0	1.7.0	1.7.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0

	emr-5.33.1	emr-5.33.0	emr-5.32.1	emr-5.32.0
Presto	0,245,1	0,245,1	0,240,1	0,240,1
Spark	2.4.7	2.4.7	2.4.7	2.4.7
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.3.1	2.3.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.9.0	0.9.0	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas da versão 5.33.1

As notas de lançamento a seguir incluem informações para a versão 5.33.0/5.33.1 do Amazon EMR. As alterações são referentes à versão 5.32.0.

Data da versão inicial: 19 de abril de 2021

Data da última atualização: 9 de agosto de 2021

Atualizações

- Atualizado o conector do Amazon Glue para a versão 1.15.0
- Atualizado para AWS SDK para Java a versão 1.11.970
- Atualizado o EMRFS para a versão 2.46.0
- Atualizado o EMR Goodies para a versão 2.14.0
- Atualizado o servidor de registros do EMR para a versão 1.9.0
- Atualizado o EMR S3 Dist CP para a versão 2.18.0
- Atualizado o agente secreto do EMR para a versão 1.8.0
- Atualizado o Flink para a versão 1.12.1

- Atualizado o Hadoop para a versão 2.10.1-amzn-1
- Atualizado o Hive para a versão 2.3.7-amzn-4
- Atualizado o Hudi para a versão 0.7.0
- Atualizado o Hue para a versão 4.9.0
- Atualizado o OpenCV para a versão 4.5.0
- Atualizado o Presto para a versão 0.245.1-amzn-0
- Atualizado o R para a versão 4.0.2
- Atualizado o Spark para a versão 2.4.7-amnz-1
- Atualizado TensorFlow para a versão 2.4.1
- Atualizado o Zeppelin para a versão 0.9.0

Alterações, melhorias e problemas resolvidos

- Esta é uma versão para corrigir problemas com o Amazon EMR Scaling quando ele não consegue up/scale reduzir a escala de um cluster com sucesso ou causa falhas no aplicativo.
- Corrigido um problema em que as solicitações de escalabilidade falhavam em um cluster grande e altamente utilizado quando os daemons do Amazon EMR no cluster estavam executando atividades de verificação de integridade, como a coleta do estado do nó do YARN e o estado do nó do HDFS. Isso estava acontecendo porque os daemons no cluster não conseguiam comunicar os dados do status de integridade de um nó aos componentes internos do Amazon EMR.
- Aprimorados os daemons do EMR no cluster para rastrear corretamente os estados dos nós quando são reutilizados endereços IP para melhorar a confiabilidade durante operações de escalabilidade.
- SPARK-29683. Corrigido um problema em que ocorriam falhas de trabalho durante a redução da escala verticalmente do cluster, pois o Spark presumia que todos os nós disponíveis estavam na lista de negação.
- YARN-9011. Corrigido um problema em que ocorriam falhas de trabalho devido a uma condição de corrida na desativação do YARN quando o cluster tentava aumentar ou reduzir a escala verticalmente.
- Corrigido problema com falhas de etapas ou tarefas durante a escalabilidade do cluster ao ser garantido que os estados dos nós fossem sempre consistentes entre os daemons do Amazon EMR no cluster e o YARN/HDFS.

- Corrigido um problema em que operações de cluster, como redução de escala verticalmente e envio de etapas, falhavam para clusters do Amazon EMR habilitados com a autenticação Kerberos. Isso ocorreu porque o daemon no cluster do Amazon EMR não renovou o tíquete Kerberos, que é necessário para se comunicar com segurança com a execução no nó primário. HDFS/YARN
- As versões mais recentes do Amazon EMR corrigem o problema com um limite menor de “Máximo de arquivos abertos” em relação às versões mais antigas AL2 no Amazon EMR. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR agora incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”.
- Configurar um cluster para corrigir problemas de desempenho do servidor de linha do tempo do Apache YARN versões 1 e 1.5

  As versões 1 e 1.5 do servidor de linha do tempo do Apache YARN podem causar problemas de performance com clusters do EMR muito ativos e grandes, especialmente com `yarn.resourcemanager.system-metrics-publisher.enabled=true`, que é a configuração padrão no Amazon EMR. Um servidor de linha do tempo do YARN v2 de código aberto resolve o problema de performance relacionado à escalabilidade do servidor de linha do tempo do YARN.

  Outras soluções alternativas para esse problema incluem:

  - Configurando yarn.resourcemanager. system-metrics-publisher.enabled=false em yarn-site.xml.
  - Habilitar a correção para esse problema na criação de um cluster, conforme descrito abaixo.

  As seguintes versões do Amazon EMR contêm uma correção para esse problema de performance do servidor de linha do tempo do YARN.

  EMR 5.30.2, 5.31.1, 5.32.1, 5.33.1, 5.34.x, 6.0.1, 6.1.1, 6.2.1, 6.3.1, 6.4.x

  Para habilitar a correção em qualquer uma das versões do Amazon EMR especificadas acima, defina essas propriedades como `true` em um arquivo JSON de configurações que é passado usando o parâmetro de comando `aws emr create-cluster`: `--configurations file://./configurations.json`. Ou habilite a correção usando a interface do usuário do console de reconfiguração.

  Exemplo de conteúdo do arquivo configurations.json:

```
[
```

```
{
"Classification": "yarn-site",
"Properties": {
"yarn.resourcemanager.system-metrics-publisher.timeline-server-v1.enable-batch":
 "true",
"yarn.resourcemanager.system-metrics-publisher.enabled": "true"
},
"Configurations": []
}
]
```

- O runtime do Spark agora é mais rápido na busca de localizações de partições do Hive Metastore para consultas de inserção do Spark.
- Atualizadas versões do componente. Para obter uma lista das versões do componente, consulte Sobre versões do Amazon EMR neste guia.
- Instalou o AWS Java SDK Bundle em cada novo cluster. Esse é um único jar contendo todos os serviços SDKs e suas dependências, em vez de jars de componentes individuais. Para obter mais informações, consulte Dependência no pacote Java SDK.
- Corrigidos os problemas de ajuste de escala gerenciado em versões anteriores do Amazon EMR e feitas melhorias para que os índices de falha de aplicações fossem significativamente reduzidos.
- O HTTPS agora está habilitado por padrão para repositórios do Amazon Linux. Se você estiver usando uma política de VPCE do Amazon S3 para restringir o acesso a buckets específicos, deverá adicionar o novo ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` do bucket do Amazon Linux à sua política (substitua `$region` pela região em que o endpoint está situado). Para obter mais informações, consulte esse tópico nos fóruns de AWS discussão. Anúncio: o Amazon Linux 2 agora oferece suporte à capacidade de usar HTTPS ao se conectar a repositórios de pacotes .

Novos recursos

- O Amazon EMR é compatível com Pontos de Acesso Amazon S3, atributo do amazon S3 que permite que você gerencie facilmente o acesso a data lakes compartilhados. Usando o alias do Ponto de Acesso Amazon S3, você pode simplificar o acesso aos dados em grande escala no Amazon EMR. Você pode usar os pontos de acesso do Amazon S3 com todas as versões do Amazon EMR sem custo adicional em todas as regiões onde o AWS Amazon EMR está disponível. Para saber mais sobre Pontos de Acesso Amazon S3 e aliases de ponto de acesso, consulte Usar um alias em estilo de bucket para seu ponto de acesso no Guia do usuário do Amazon S3.

- O Amazon EMR-5.33 suporta novos tipos de EC2 instância da Amazon: c5a, c5ad, c6gn, c6gd, m6gd, d3, d3en, m5zn, r5b, r6gd. Consulte Tipos de instância compatíveis.

Problemas conhecidos

- Limite inferior de “Máximo de arquivos abertos” em versões mais antigas AL2 [corrigido em versões mais recentes]. Versões do Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 e emr-6.2.0 são baseadas em versões mais antigas do Amazon Linux 2 (), que AL2 têm uma configuração de limite inferior para “Máximo de arquivos abertos” quando clusters do Amazon EMR são criados com a AMI padrão. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”. Versões com o limite inferior de arquivos abertos causam o erro “Muitos arquivos abertos” ao ser enviado um trabalho do Spark. Nas versões afetadas, a AMI padrão do Amazon EMR tem uma configuração de ulimit padrão de 4096 para “Máximo de arquivos abertos”, que é inferior ao limite de 65536 arquivos na AMI mais recente do Amazon Linux 2. A configuração inferior de ulimit para “Máximo de arquivos abertos” causa falhas em trabalhos do Spark quando o driver e o executor do Spark tentam abrir mais de 4096 arquivos. Para corrigir o problema, o Amazon EMR tem um script de ação de bootstrap (BA) que ajusta a configuração de ulimit na criação do cluster.

  Se você está usando uma versão mais antiga do Amazon EMR que não tem a correção permanente para esse problema, a solução alternativa a seguir permite que você defina explicitamente o ulimit instance-controller para um máximo de 65536 arquivos.

  Defina explicitamente um ulimit na linha de comando

  1. Edite `/etc/systemd/system/instance-controller.service` para adicionar os seguintes parâmetros à seção Serviço.

     ```
     LimitNOFILE=65536

     LimitNPROC=65536
     ```

  2. Reiniciar InstanceController

     ```
     $ sudo systemctl daemon-reload

     $ sudo systemctl restart instance-controller
     ```

Defina um ulimit usando a ação de bootstrap (BA)

Você também pode usar um script de ação de bootstrap (BA) para configurar o ulimit instance-controller para 65536 arquivos na criação do cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- Para clusters de sub-rede privados do Amazon EMR 6.3.0 e 6.2.0, você não pode acessar a interface do usuário da Web do Ganglia. Você receberá um erro de “acesso negado (403)”. Outros sites UIs, como Spark, Hue, Zeppelin JupyterHub, Livy e Tez, estão funcionando normalmente. O acesso à interface do usuário da Web do Ganglia em clusters de sub-redes públicas também está funcionando normalmente. Para resolver esse problema, reinicie o serviço httpd no nó primário com `sudo systemctl restart httpd`. Esse problema foi corrigido na versão 6.4.0 do Amazon EMR.
- **Important**

  Os clusters do EMR que executam Amazon Linux ou Amazon Linux 2 Amazon Machine Images (AMIs) usam o comportamento padrão do Amazon Linux e não baixam e instalam automaticamente atualizações importantes e críticas do kernel que exigem uma reinicialização. Esse é o mesmo comportamento de outras EC2 instâncias da Amazon que executam o Amazon Linux AMI padrão. Se novas atualizações de software do Amazon Linux que exigem reinicialização (como atualizações do kernel, NVIDIA e CUDA) forem

disponibilizadas após o lançamento de uma versão do Amazon EMR, as instâncias de cluster do Amazon EMR que executam a AMI padrão não baixarão nem instalarão essas atualizações automaticamente. Para obter atualizações do kernel, você pode personalizar sua AMI do Amazon EMR para usar a AMI do Amazon Linux mais recente.

- O suporte do console para criar uma configuração de segurança que especifica a opção de integração do AWS Ranger atualmente não é suportado na GovCloud região. A configuração de segurança do pode ser feita usando a CLI. Consulte Criar a configuração de segurança do EMR no Guia de gerenciamento do Amazon EMR.
- Políticas gerenciadas com escopo definido: para se alinhar às AWS melhores práticas, o Amazon EMR introduziu políticas gerenciadas padrão com escopo do EMR v2 como substitutas das políticas que serão descontinuadas. Consulte as políticas gerenciadas do Amazon EMR.
- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
  - O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

  Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.

Versões de componente 5.33.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a

versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *CommunityVersion*-amzn-*EmrVersion*. O *EmrVersion* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.1	SDK do Amazon SageMaker Spark
emr-ddb	4.16.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.14.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.2.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.18.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.6.0	EMR S3Select Connector
emrfs	2.46.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.

Componente	Versão	Descrição
flink-client	1.12.1	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.12.1	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.10.1-amzn-1.1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.10.1-amzn-1.1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.10.1-amzn-1.1	O cliente de linha de comando e biblioteca do HDFS

Componente	Versão	Descrição
hadoop-hdfs-namenode	2.10.1-amzn-1.1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	2.10.1-amzn-1.1	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	2.10.1-amzn-1.1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.10.1-amzn-1.1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.10.1-amzn-1.1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.10.1-amzn-1.1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.10.1-amzn-1.1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.10.1-amzn-1.1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.

Componente	Versão	Descrição
hbase-hmaster	1.4.13	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.13	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.13	HBase cliente de linha de comando.
hbase-rest-server	1.4.13	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.13	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.7-amzn-4	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.7-amzn-4	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.3.7-amzn-4	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.7-amzn-4	O cliente da linha de comando do Hive.
hive-hbase	2.3.7-amzn-4	Cliente do Hive-hbase.

Componente	Versão	Descrição
hive-metastore-server	2.3.7-amzn-4	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.7-amzn-4	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.7.0-amzn-1	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-spark	0.7.0-amzn-1	Biblioteca de pacotes para executar o Spark com o Hudi.
hudi-presto	0.7.0-amzn-1	Biblioteca de pacotes para executar o Presto com o Hudi.
hue-server	4.9.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	1.2.2	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP

Componente	Versão	Descrição
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.7.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68+	O servidor de banco de dados MySQL.
nvidia-cuda	10.1.243	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.0	O cliente da linha de comando do Oozie.
oozie-server	5.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.5.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.3- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.3- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.245.1-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.

Componente	Versão	Descrição
presto-worker	0.245.1-amzn-0	O serviço que executa partes de uma consulta.
presto-client	0.245.1-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	4.0.2	O projeto R para computação estatística
ranger-kms-server	1.2.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	2.4.7-amzn-1.1	Os clientes da linha de comando do Spark.
spark-history-server	2.4.7-amzn-1.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.7-amzn-1.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.4.7-amzn-1.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.

Componente	Versão	Descrição
tensorflow	2.4.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.9.0	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.33.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-5.33.1

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

Classificações	Descrição	Ações de reconfiguração
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Not available.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.

Classificações	Descrição	Ações de reconfiguração
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.

Classificações	Descrição	Ações de reconfiguração
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	Should not be reconfigured.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

Classificações	Descrição	Ações de reconfiguração
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.pr operties da HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Restarts HiveServer2 and HiveMetastore.

Classificações	Descrição	Ações de reconfiguração
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j	Alterar as configurações de log4j.properties no Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server.

Classificações	Descrição	Ações de reconfiguração
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server.
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
recordserver-env	Altere os valores no ambiente do EMR. RecordServer	Restarts EMR record server.
recordserver-conf	Altere os valores no arquivo erver.properties RecordServer do EMR.	Restarts EMR record server.
recordserver-log4j	Altere os valores no arquivo log4j.properties RecordServer do EMR.	Restarts EMR record server.
spark	As configurações administr adas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.	Restarts Spark history server and Spark thrift server.

Classificações	Descrição	Ações de reconfiguração
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restarts Oozie.
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.

Classificações	Descrição	Ações de reconfiguração
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Versão 5.33.0 do Amazon EMR

Versões 5.33.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterEnterpriseGateway, JupyterHubLivy, MXNet, Mahout, Oozie, PhoenixPig, Presto, Spark, SqoopTensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.33.0	emr-5.32.1	emr-5.32.0	emr-5.31.1
AWS SDK for Java	1.11.970	1.11.890	1.11.890	1.11.852
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12

	emr-5.33.0	emr-5.32.1	emr-5.32.0	emr-5.31.1
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.12.1	1.11.2	1.11.2	1.11.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.7	2.3.7	2.3.7	2.3.7
Hadoop	2.10.1	2.10.1	2.10.1	2.10.0
Hive	2.3.7	2.3.7	2.3.7	2.3.7
Hudi	0.7.0-amzn-1	0.6.0-amzn-0	0.6.0-amzn-0	0.6.0-amzn-0
Hue	4.9.0	4.8.0	4.8.0	4.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	-
JupyterHub	1.2.2	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.7.0	1.7.0	1.6.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0

	emr-5.33.0	emr-5.32.1	emr-5.32.0	emr-5.31.1
Presto	0,245,1	0,240,1	0,240,1	0,238,3
Spark	2.4.7	2.4.7	2.4.7	2.4.6
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.3.1	2.3.1	2.1.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.9.0	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas da versão 5.33.0

Versões de componente 5.33.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.1	SDK do Amazon SageMaker Spark
emr-ddb	4.16.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.14.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.2.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.18.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.6.0	EMR S3Select Connector
emrfs	2.46.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.12.1	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.12.1	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager

Componente	Versão	Descrição
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.10.1-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.10.1-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.10.1-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.10.1-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	2.10.1-amzn-1	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.

Componente	Versão	Descrição
hadoop-httpfs-server	2.10.1-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.10.1-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.10.1-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.10.1-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.10.1-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.10.1-amzn-1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.4.13	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.13	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.13	HBase cliente de linha de comando.

Componente	Versão	Descrição
hbase-rest-server	1.4.13	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.13	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.7-amzn-4	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.7-amzn-4	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.3.7-amzn-4	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.7-amzn-4	O cliente da linha de comando do Hive.
hive-hbase	2.3.7-amzn-4	Cliente do Hive-hbase.
hive-metastore-server	2.3.7-amzn-4	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.7-amzn-4	O serviço que aceita as consultas do Hive como solicitações da web.

Componente	Versão	Descrição
hudi	0.7.0-amzn-1	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-spark	0.7.0-amzn-1	Biblioteca de pacotes para executar o Spark com o Hudi.
hudi-presto	0.7.0-amzn-1	Biblioteca de pacotes para executar o Presto com o Hudi.
hue-server	4.9.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	1.2.2	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.7.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68	O servidor de banco de dados MySQL.
nvidia-cuda	10.1.243	Drivers NVIDIA e toolkit CUDA

Componente	Versão	Descrição
oozie-client	5.2.0	O cliente da linha de comando do Oozie.
oozie-server	5.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.5.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.3- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.3- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.245.1-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.245.1-amzn-0	O serviço que executa partes de uma consulta.
presto-client	0.245.1-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.

Componente	Versão	Descrição
r	4.0.2	O projeto R para computação estatística
ranger-kms-server	1.2.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	2.4.7-amzn-1	Os clientes da linha de comando do Spark.
spark-history-server	2.4.7-amzn-1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.7-amzn-1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.4.7-amzn-1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.4.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.

Componente	Versão	Descrição
zeppelin-server	0.9.0	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.33.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-5.33.0

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.

Classificações	Descrição	Ações de reconfiguração
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-session	Altere as configurações do Flink log4j-session.properties para a sessão. Kubernetes/ Yarn	Not available.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.

Classificações	Descrição	Ações de reconfiguração
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.

Classificações	Descrição	Ações de reconfiguração
hbase	Configurações seleciona das pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hb ase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	Should not be reconfigured.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

Classificações	Descrição	Ações de reconfiguração
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.pr operties da HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Restarts HiveServer2 and HiveMetastore.

Classificações	Descrição	Ações de reconfiguração
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j	Alterar as configurações de log4j.properties no Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server.

Classificações	Descrição	Ações de reconfiguração
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server.
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
recordserver-env	Altere os valores no ambiente do EMR. RecordServer	Restarts EMR record server.
recordserver-conf	Altere os valores no arquivo erver.properties RecordServer do EMR.	Restarts EMR record server.
recordserver-log4j	Altere os valores no arquivo log4j.properties RecordServer do EMR.	Restarts EMR record server.
spark	As configurações administr adas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.	Restarts Spark history server and Spark thrift server.

Classificações	Descrição	Ações de reconfiguração
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restarts Oozie.
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zeppelin-site	Alterar as definições da configuração em zeppelin-site.xml.	Restarts Zeppelin.

Classificações	Descrição	Ações de reconfiguração
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Versão 5.32.1 do Amazon EMR

Versões 5.32.1 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterEnterpriseGateway, JupyterHubLivy, MXNet, Mahout, Oozie, PhoenixPig, Presto, Spark, SqoopTensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.32.1	emr-5.32.0	emr-5.31.1	emr-5.31.0
AWS SDK for Java	1.11.890	1.11.890	1.11.852	1.11.852
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12

	emr-5.32.1	emr-5.32.0	emr-5.31.1	emr-5.31.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.2	1.11.2	1.11.0	1.11.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.7	2.3.7	2.3.7	2.3.7
Hadoop	2.10.1	2.10.1	2.10.0	2.10.0
Hive	2.3.7	2.3.7	2.3.7	2.3.7
Hudi	0.6.0-amzn-0	0.6.0-amzn-0	0.6.0-amzn-0	0.6.0-amzn-0
Hue	4.8.0	4.8.0	4.7.1	4.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.7.0	1.6.0	1.6.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0

	emr-5.32.1	emr-5.32.0	emr-5.31.1	emr-5.31.0
Presto	0,240,1	0,240,1	0,238,3	0,238,3
Spark	2.4.7	2.4.7	2.4.6	2.4.6
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.3.1	2.3.1	2.1.0	2.1.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas da versão 5.32.1

Esta é uma versão para corrigir problemas com o Amazon EMR Scaling quando ele não consegue up/scale reduzir a escala de um cluster com sucesso ou causa falhas no aplicativo.

Alterações, melhorias e problemas resolvidos

- Corrigido um problema em que as solicitações de escalabilidade falhavam em um cluster grande e altamente utilizado quando os daemons do Amazon EMR no cluster estavam executando atividades de verificação de integridade, como a coleta do estado do nó do YARN e o estado do nó do HDFS. Isso estava acontecendo porque os daemons no cluster não conseguiam comunicar os dados do status de integridade de um nó aos componentes internos do Amazon EMR.
- Aprimorados os daemons do EMR no cluster para rastrear corretamente os estados dos nós quando são reutilizados endereços IP para melhorar a confiabilidade durante operações de escalabilidade.
- SPARK-29683. Corrigido um problema em que ocorriam falhas de trabalho durante a redução da escala verticalmente do cluster, pois o Spark presumia que todos os nós disponíveis estavam na lista de negação.

- YARN-9011. Corrigido um problema em que ocorriam falhas de trabalho devido a uma condição de corrida na desativação do YARN quando o cluster tentava aumentar ou reduzir a escala verticalmente.
- Corrigido problema com falhas de etapas ou tarefas durante a escalabilidade do cluster ao ser garantido que os estados dos nós fossem sempre consistentes entre os daemons do Amazon EMR no cluster e o YARN/HDFS.
- Corrigido um problema em que operações de cluster, como redução de escala verticalmente e envio de etapas, falhavam para clusters do Amazon EMR habilitados com a autenticação Kerberos. Isso ocorreu porque o daemon no cluster do Amazon EMR não renovou o tíquete Kerberos, que é necessário para se comunicar com segurança com a execução no nó primário. HDFS/YARN
- As versões mais recentes do Amazon EMR corrigem o problema com um limite menor de “Máximo de arquivos abertos” em relação às versões mais antigas AL2 no Amazon EMR. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR agora incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”.
- O HTTPS agora está habilitado por padrão para repositórios do Amazon Linux. Se você estiver usando uma política de VPCE do Amazon S3 para restringir o acesso a buckets específicos, deverá adicionar o novo ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` do bucket do Amazon Linux à sua política (substitua `$region` pela região em que o endpoint está situado). Para obter mais informações, consulte esse tópico nos fóruns de AWS discussão. Anúncio: o Amazon Linux 2 agora oferece suporte à capacidade de usar HTTPS ao se conectar a repositórios de pacotes .

Problemas conhecidos

- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.

- O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere / (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.

Versões de componente 5.32.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.1	SDK do Amazon SageMaker Spark
emr-ddb	4.16.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.13.0	Bibliotecas convenientes para o ecossistema do Hadoop.

Componente	Versão	Descrição
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.1.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.17.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.6.0	EMR S3Select Connector
emrfs	2.45.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.11.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.11.2	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.10.1-amzn-0.1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.10.1-amzn-0.1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.10.1-amzn-0.1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.10.1-amzn-0.1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	2.10.1-amzn-0.1	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	2.10.1-amzn-0.1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.10.1-amzn-0.1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.10.1-amzn-0.1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.

Componente	Versão	Descrição
hadoop-yarn-nodemanager	2.10.1-amzn-0.1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.10.1-amzn-0.1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hadoop-yarn-timeline-server	2.10.1-amzn-0.1	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.4.13	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.13	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.13	HBase cliente de linha de comando.
hbase-rest-server	1.4.13	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.13	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.7-amzn-3	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.

Componente	Versão	Descrição
hcatalog-server	2.3.7-amzn-3	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.3.7-amzn-3	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.7-amzn-3	O cliente da linha de comando do Hive.
hive-hbase	2.3.7-amzn-3	Cliente do Hive-hbase.
hive-metastore-server	2.3.7-amzn-3	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.7-amzn-3	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.6.0-amzn-0	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-spark	0.6.0-amzn-0	Biblioteca de pacotes para executar o Spark com o Hudi.
hudi-presto	0.6.0-amzn-0	Biblioteca de pacotes para executar o Presto com o Hudi.

Componente	Versão	Descrição
hue-server	4.8.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	1.1.0	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.7.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68+	O servidor de banco de dados MySQL.
nvidia-cuda	10.1.243	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.0	O cliente da linha de comando do Oozie.
oozie-server	5.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.4.0	Biblioteca de Visão Computaci onal de Código Aberto.

Componente	Versão	Descrição
phoenix-library	4.14.3- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.3- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.240.1-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.240.1-amzn-0	O serviço que executa partes de uma consulta.
presto-client	0.240.1-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.3	O projeto R para computação estatística
ranger-kms-server	1.2.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	2.4.7-amzn-0.1	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	2.4.7-amzn-0.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.7-amzn-0.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.4.7-amzn-0.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.3.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.8.2	O notebook baseado na web que permite a análise de dados interativa.
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.

Componente	Versão	Descrição
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.32.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-5.32.1

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceM anager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager,

Classificações	Descrição	Ações de reconfiguração
		NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce- HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.	Not available.

Classificações	Descrição	Ações de reconfiguração
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.

Classificações	Descrição	Ações de reconfiguração
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hbase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	Should not be reconfigured.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

Classificações	Descrição	Ações de reconfiguração
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.pr operties da HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Restarts HiveServer2 and HiveMetastore.

Classificações	Descrição	Ações de reconfiguração
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j	Alterar as configurações de log4j.properties no Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server.

Classificações	Descrição	Ações de reconfiguração
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server.
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
recordserver-env	Altere os valores no ambiente do EMR. RecordServer	Restarts EMR record server.
recordserver-conf	Altere os valores no arquivo erver.properties RecordServer do EMR.	Restarts EMR record server.
recordserver-log4j	Altere os valores no arquivo log4j.properties RecordServer do EMR.	Restarts EMR record server.
spark	As configurações administr adas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.	Restarts Spark history server and Spark thrift server.

Classificações	Descrição	Ações de reconfiguração
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restarts Oozie.
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.

Classificações	Descrição	Ações de reconfiguração
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Versão 5.32.0 do Amazon EMR

Versões 5.32.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterEnterpriseGateway, JupyterHubLivy, MXNet, Mahout, Oozie, PhoenixPig, Presto, Spark, SqoopTensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.32.0	emr-5.31.1	emr-5.31.0	emr-5.30.2
AWS SDK for Java	1.11.890	1.11.852	1.11.852	1.11.759
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-

	emr-5.32.0	emr-5.31.1	emr-5.31.0	emr-5.30.2
Delta	-	-	-	-
Flink	1.11.2	1.11.0	1.11.0	1.10.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.7	2.3.7	2.3.7	2.3.6
Hadoop	2.10.1	2.10.0	2.10.0	2.8.5
Hive	2.3.7	2.3.7	2.3.7	2.3.6
Hudi	0.6.0-amzn-0	0.6.0-amzn-0	0.6.0-amzn-0	0.5.2-incubating
Hue	4.8.0	4.7.1	4.7.1	4.6.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	-	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.6.0	1.6.0	1.5.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,240,1	0,238,3	0,238,3	0.232
Spark	2.4.7	2.4.6	2.4.6	2.4.5

	emr-5.32.0	emr-5.31.1	emr-5.31.0	emr-5.30.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.3.1	2.1.0	2.1.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas da versão 5.32.0

As notas da versão a seguir incluem informações para a versão 5.32.0 do Amazon EMR. As alterações são referentes à versão 5.31.0.

Data da versão inicial: 8 de janeiro de 2021

Atualizações

- Atualizado o conector do Amazon Glue para a versão 1.14.0
- SDK do Amazon SageMaker Spark atualizado para a versão 1.4.1
- Atualizado para AWS SDK para Java a versão 1.11.890
- Atualizado o conector do DynamoDB para EMR, versão 4.16.0
- Atualizado o EMRFS para a versão 2.45.0
- Atualizadas as métricas de analytics de log do EMR para a versão 1.18.0
- Cliente MetricsAndEventsApiGateway EMR atualizado para a versão 1.5.0
- Atualizado o servidor de registros do EMR para a versão 1.8.0
- Atualizado o EMR S3 Dist CP para a versão 2.17.0
- Atualizado o agente secreto do EMR para a versão 1.7.0
- Atualizado o Flink para a versão 1.11.2
- Atualizado o Hadoop para a versão 2.10.1-amzn-0

- Atualizado o Hive para a versão 2.3.7-amzn-3
- Atualizado o Hue para a versão 4.8.0
- Atualizado o Mxnet para a versão 1.7.0
- Atualizado o OpenCV para a versão 4.4.0
- Atualizado o Presto para a versão 0.240.1-amzn-0
- Atualizado o Spark para a versão 2.4.7-amnz-0
- Atualizado TensorFlow para a versão 2.3.1

Alterações, melhorias e problemas resolvidos

- Esta é uma versão para corrigir problemas com o Amazon EMR Scaling quando ele não consegue up/scale reduzir a escala de um cluster com sucesso ou causa falhas no aplicativo.
- Corrigido um problema em que as solicitações de escalabilidade falhavam em um cluster grande e altamente utilizado quando os daemons do Amazon EMR no cluster estavam executando atividades de verificação de integridade, como a coleta do estado do nó do YARN e o estado do nó do HDFS. Isso estava acontecendo porque os daemons no cluster não conseguiam comunicar os dados do status de integridade de um nó aos componentes internos do Amazon EMR.
- Aprimorados os daemons do EMR no cluster para rastrear corretamente os estados dos nós quando são reutilizados endereços IP para melhorar a confiabilidade durante operações de escalabilidade.
- SPARK-29683. Corrigido um problema em que ocorriam falhas de trabalho durante a redução da escala verticalmente do cluster, pois o Spark presumia que todos os nós disponíveis estavam na lista de negação.
- YARN-9011. Corrigido um problema em que ocorriam falhas de trabalho devido a uma condição de corrida na desativação do YARN quando o cluster tentava aumentar ou reduzir a escala verticalmente.
- Corrigido problema com falhas de etapas ou tarefas durante a escalabilidade do cluster ao ser garantido que os estados dos nós fossem sempre consistentes entre os daemons do Amazon EMR no cluster e o YARN/HDFS.
- Corrigido um problema em que operações de cluster, como redução de escala verticalmente e envio de etapas, falhavam para clusters do Amazon EMR habilitados com a autenticação Kerberos. Isso ocorreu porque o daemon no cluster do Amazon EMR não renovou o tíquete Kerberos, que é necessário para se comunicar com segurança com a execução no nó primário. HDFS/YARN

- As versões mais recentes do Amazon EMR corrigem o problema com um limite menor de “Máximo de arquivos abertos” em relação às versões mais antigas AL2 no Amazon EMR. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR agora incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”.
- Atualizadas versões do componente.
- Para obter uma lista das versões do componente, consulte Sobre versões do Amazon EMR neste guia.

Novos recursos

- Desde as versões 5.32.0 e 6.5.0 do Amazon EMR, o dimensionamento do executor dinâmico para o Apache Spark está habilitado por padrão. Para ativar ou desativar esse atributo, você pode usar o parâmetro de configuração `spark.yarn.heterogeneousExecutors.enabled`.
- Status de suporte do Instance Metadata Service (IMDS) V2: os componentes do Amazon EMR 5.23.1, 5.27.1 e 5.32 ou posteriores são usados para todas as chamadas do IMDS. IMDSv2 Para chamadas IMDS no código do aplicativo, você pode usar ambos IMDSv1 e IMDSv2, ou configurar o IMDS para uso somente IMDSv2 para aumentar a segurança. Para outras versões 5.x do EMR, a IMDSv1 desativação causa falha na inicialização do cluster.
- Desde a versão 5.32.0 do Amazon EMR, você pode iniciar um cluster que se integre nativamente ao Apache Ranger. O Apache Ranger é uma estrutura de código aberto para habilitar, monitorar e gerenciar uma segurança de dados abrangente em toda a plataforma Hadoop. Para obter mais informações, consulte Apache Ranger. Com a integração nativa, você pode trazer seu próprio Apache Ranger para aplicar um controle de acesso detalhado aos dados no Amazon EMR. Consulte Integrar o Amazon EMR com o Apache Ranger no Guia de lançamento do Amazon EMR.
- A versão 5.32.0 do Amazon EMR é compatível com o Amazon EMR no EKS. Para obter mais detalhes sobre os conceitos básicos do EMR no EKS, consulte O que é o Amazon EMR no EKS.
- A versão 5.32.0 do Amazon EMR é compatível com o Amazon EMR Studio (versão de demonstração). Para obter mais detalhes sobre os conceitos básicos do EMR Studio, consulte Amazon EMR Studio (versão de demonstração).
- Políticas gerenciadas com escopo definido: para se alinhar às AWS melhores práticas, o Amazon EMR introduziu políticas gerenciadas padrão com escopo do EMR v2 como substitutas das políticas que serão descontinuadas. Consulte as políticas gerenciadas do Amazon EMR.

Problemas conhecidos

- Para clusters de sub-rede privados do Amazon EMR 6.3.0 e 6.2.0, você não pode acessar a interface do usuário da Web do Ganglia. Você receberá um erro de “acesso negado (403)”. Outros sites UIs, como Spark, Hue, Zeppelin JupyterHub, Livy e Tez, estão funcionando normalmente. O acesso à interface do usuário da Web do Ganglia em clusters de sub-redes públicas também está funcionando normalmente. Para resolver esse problema, reinicie o serviço httpd no nó primário com `sudo systemctl restart httpd`. Esse problema foi corrigido na versão 6.4.0 do Amazon EMR.
- Limite inferior de “Máximo de arquivos abertos” em versões mais antigas AL2 [corrigido em versões mais recentes]. Versões do Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 e emr-6.2.0 são baseadas em versões mais antigas do Amazon Linux 2 (), que AL2 têm uma configuração de limite inferior para “Máximo de arquivos abertos” quando clusters do Amazon EMR são criados com a AMI padrão. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”. Versões com o limite inferior de arquivos abertos causam o erro “Muitos arquivos abertos” ao ser enviado um trabalho do Spark. Nas versões afetadas, a AMI padrão do Amazon EMR tem uma configuração de ulimit padrão de 4096 para “Máximo de arquivos abertos”, que é inferior ao limite de 65536 arquivos na AMI mais recente do Amazon Linux 2. A configuração inferior de ulimit para “Máximo de arquivos abertos” causa falhas em trabalhos do Spark quando o driver e o executor do Spark tentam abrir mais de 4096 arquivos. Para corrigir o problema, o Amazon EMR tem um script de ação de bootstrap (BA) que ajusta a configuração de ulimit na criação do cluster.

  Se você está usando uma versão mais antiga do Amazon EMR que não tem a correção permanente para esse problema, a solução alternativa a seguir permite que você defina explicitamente o ulimit instance-controller para um máximo de 65536 arquivos.

  Defina explicitamente um ulimit na linha de comando

  1. Edite `/etc/systemd/system/instance-controller.service` para adicionar os seguintes parâmetros à seção Serviço.

     ```
     LimitNOFILE=65536

     LimitNPROC=65536
     ```

  2. Reiniciar InstanceController

```
$ sudo systemctl daemon-reload

$ sudo systemctl restart instance-controller
```

Defina um ulimit usando a ação de bootstrap (BA)

Você também pode usar um script de ação de bootstrap (BA) para configurar o ulimit instance-controller para 65536 arquivos na criação do cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- **Important**

  Os clusters do EMR que executam Amazon Linux ou Amazon Linux 2 Amazon Machine Images (AMIs) usam o comportamento padrão do Amazon Linux e não baixam e instalam automaticamente atualizações importantes e críticas do kernel que exigem uma reinicialização. Esse é o mesmo comportamento de outras EC2 instâncias da Amazon que executam o Amazon Linux AMI padrão. Se novas atualizações de software do Amazon Linux que exigem reinicialização (como atualizações do kernel, NVIDIA e CUDA) forem disponibilizadas após o lançamento de uma versão do Amazon EMR, as instâncias de cluster do Amazon EMR que executam a AMI padrão não baixarão nem instalarão essas

atualizações automaticamente. Para obter atualizações do kernel, você pode personalizar sua AMI do Amazon EMR para usar a AMI do Amazon Linux mais recente.

- O suporte do console para criar uma configuração de segurança que especifica a opção de integração do AWS Ranger atualmente não é suportado na GovCloud região. A configuração de segurança do pode ser feita usando a CLI. Consulte Criar a configuração de segurança do EMR no Guia de gerenciamento do Amazon EMR.
- Quando AtRestEncryption a criptografia HDFS é habilitada em um cluster que usa o Amazon EMR 5.31.0 ou 5.32.0, as consultas do Hive resultam na seguinte exceção de tempo de execução.

```
TaskAttempt 3 failed, info=[Error: Error while running task ( failure ) :
 attempt_1604112648850_0001_1_01_000000_3:java.lang.RuntimeException:
 java.lang.RuntimeException: Hive Runtime Error while closing
 operators: java.io.IOException: java.util.ServiceConfigurationError:
 org.apache.hadoop.security.token.TokenIdentifier: Provider
 org.apache.hadoop.hbase.security.token.AuthenticationTokenIdentifier not found
```

- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
  - O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

  Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.

Versões de componente 5.32.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.1	SDK do Amazon SageMaker Spark
emr-ddb	4.16.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.13.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.1.0	Ambiente Conda para bloco de anotações do emr, que inclui o jupyter enterprise gateway
emr-s3-dist-cp	2.17.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.

Componente	Versão	Descrição
emr-s3-select	1.6.0	EMR S3Select Connector
emrfs	2.45.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.11.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.11.2	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.10.1-amzn-0	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.10.1-amzn-0	O serviço de nível de nó do HDFS para armazenamento de blocos.

Componente	Versão	Descrição
hadoop-hdfs-library	2.10.1-amzn-0	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.10.1-amzn-0	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	2.10.1-amzn-0	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	2.10.1-amzn-0	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.10.1-amzn-0	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.10.1-amzn-0	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.10.1-amzn-0	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.10.1-amzn-0	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hadoop-yarn-timeline-server	2.10.1-amzn-0	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.

Componente	Versão	Descrição
hbase-hmaster	1.4.13	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.13	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.13	HBase cliente de linha de comando.
hbase-rest-server	1.4.13	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.13	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.7-amzn-3	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.7-amzn-3	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.3.7-amzn-3	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.7-amzn-3	O cliente da linha de comando do Hive.
hive-hbase	2.3.7-amzn-3	Cliente do Hive-hbase.

Componente	Versão	Descrição
hive-metastore-server	2.3.7-amzn-3	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.7-amzn-3	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.6.0-amzn-0	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-spark	0.6.0-amzn-0	Biblioteca de pacotes para executar o Spark com o Hudi.
hudi-presto	0.6.0-amzn-0	Biblioteca de pacotes para executar o Presto com o Hudi.
hue-server	4.8.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	1.1.0	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP

Componente	Versão	Descrição
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.7.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.68	O servidor de banco de dados MySQL.
nvidia-cuda	10.1.243	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.0	O cliente da linha de comando do Oozie.
oozie-server	5.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.4.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.3- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.3- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.240.1-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.

Componente	Versão	Descrição
presto-worker	0.240.1-amzn-0	O serviço que executa partes de uma consulta.
presto-client	0.240.1-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.3	O projeto R para computação estatística
ranger-kms-server	1.2.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	2.4.7-amzn-0	Os clientes da linha de comando do Spark.
spark-history-server	2.4.7-amzn-0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.7-amzn-0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.4.7-amzn-0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.

Componente	Versão	Descrição
tensorflow	2.3.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.8.2	O notebook baseado na web que permite a análise de dados interativa.
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.32.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Ocorrem ações de reconfiguração quando você especifica uma configuração para grupos de instâncias em um cluster em execução. O Amazon EMR só inicia ações de reconfiguração para as classificações que você modifica. Para obter mais informações, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Classificações do emr-5.32.0

Classificações	Descrição	Ações de reconfiguração
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.	Restarts the ResourceManager service.
container-executor	Alterar os valores no arquivo container-executor.cfg do YARN do Hadoop.	Not available.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.	Not available.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Alterar as configurações relacionadas ao docker.	Not available.
emrfs-site	Alterar as configurações do EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

Classificações	Descrição	Ações de reconfiguração
		Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Alterar as configurações do flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Alterar as configurações de log4j.properties no Flink.	Restarts Flink history server.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.	Not available.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.	Restarts Flink history server.

Classificações	Descrição	Ações de reconfiguração
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop	Not available.
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop	Not available.

Classificações	Descrição	Ações de reconfiguração
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Mude os valores no HBase ambiente.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hbase.properties do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.	Not available.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.	Should not be reconfigured.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

Classificações	Descrição	Ações de reconfiguração
hcatalog-env	Mude os valores no HCatalog ambiente.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.pr operties da HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.	Not available.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.	Not available.
hive-env	Alterar os valores no ambiente do Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.	Restarts HiveServer2 and HiveMetastore.

Classificações	Descrição	Ações de reconfiguração
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.	Not available.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.	Not available.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.	Not available.
hue-ini	Alterar os valores no arquivo ini do Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Alterar os valores no ambiente do HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.	Not available.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.	Not available.

Classificações	Descrição	Ações de reconfiguração
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Altere os valores no ambiente do Hudi.	Not available.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.	Not available.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.	Not available.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.	Not available.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.	Not available.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.	Restarts Livy Server.
livy-env	Alterar os valores no ambiente do Livy.	Restarts Livy Server.
livy-log4j	Alterar as configurações de log4j.properties no Livy.	Restarts Livy Server.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.	Restarts Hadoop MapReduce-HistoryServer.

Classificações	Descrição	Ações de reconfiguração
oozie-env	Alterar os valores no ambiente do Oozie.	Restarts Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.	Restarts Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.	Not available.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.	Not available.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.	Not available.
pig-env	Alterar os valores no ambiente do Pig.	Not available.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.	Restarts Oozie.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.	Not available.
presto-log	Alterar os valores no arquivo log.properties do Presto.	Restarts Presto-Server.

Classificações	Descrição	Ações de reconfiguração
presto-config	Alterar os valores no arquivo config.properties do Presto.	Restarts Presto-Server.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.	Not available.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.	Restarts Presto-Server.
presto-node	Altere os valores no arquivo node.properties do Presto.	Not available.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.	Not available.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.	Not available.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.	Restarts Presto-Server.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.	Not available.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.	Not available.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.	Not available.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.	Not available.

Classificações	Descrição	Ações de reconfiguração
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.	Not available.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.	Not available.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.	Not available.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.	Not available.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.	Not available.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.	Not available.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.	Not available.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.	Not available.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.	Restarts Ranger KMS Server.

Classificações	Descrição	Ações de reconfiguração
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.	Not available.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.	Not available.
recordserver-env	Altere os valores no ambiente do EMR. RecordServer	Restarts EMR record server.
recordserver-conf	Altere os valores no arquivo erver.properties RecordServer do EMR.	Restarts EMR record server.
recordserver-log4j	Altere os valores no arquivo log4j.properties RecordServer do EMR.	Restarts EMR record server.
spark	As configurações administr adas pelo Amazon EMR para o Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Alterar os valores no ambiente do Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.	Not available.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.	Restarts Spark history server and Spark thrift server.

Classificações	Descrição	Ações de reconfiguração
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Alterar os valores no ambiente do Sqoop.	Not available.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.	Not available.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.	Not available.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.	Restarts Oozie.
yarn-env	Alterar os valores no ambiente do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.	Restarts Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.	Restarts Zookeeper server.

Classificações	Descrição	Ações de reconfiguração
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.	Restarts Zookeeper server.

Versão 5.31.1 do Amazon EMR

Versões 5.31.1 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterHub, LivyMXNet, Mahout, Oozie, Phoenix, PigPresto, Spark, Sqoop, TensorFlow, TezZeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.31.1	emr-5.31.0	emr-5.30.2	emr-5.30.1
AWS SDK for Java	1.11.852	1.11.852	1.11.759	1.11.759
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-

	emr-5.31.1	emr-5.31.0	emr-5.30.2	emr-5.30.1
Delta	-	-	-	-
Flink	1.11.0	1.11.0	1.10.0	1.10.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.7	2.3.7	2.3.6	2.3.6
Hadoop	2.10.0	2.10.0	2.8.5	2.8.5
Hive	2.3.7	2.3.7	2.3.6	2.3.6
Hudi	0.6.0-amzn-0	0.6.0-amzn-0	0.5.2-incubating	0.5.2-incubating
Hue	4.7.1	4.7.1	4.6.0	4.6.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.6.0	1.6.0	1.5.1	1.5.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,238,3	0,238,3	0.232	0.232
Spark	2.4.6	2.4.6	2.4.5	2.4.5

	emr-5.31.1	emr-5.31.0	emr-5.30.2	emr-5.30.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.1.0	2.1.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas da versão 5.31.1

Esta é uma versão para corrigir problemas com o Amazon EMR Scaling quando ele não consegue up/scale reduzir a escala de um cluster com sucesso ou causa falhas no aplicativo.

Alterações, melhorias e problemas resolvidos

- Corrigido um problema em que as solicitações de escalabilidade falhavam em um cluster grande e altamente utilizado quando os daemons do Amazon EMR no cluster estavam executando atividades de verificação de integridade, como a coleta do estado do nó do YARN e o estado do nó do HDFS. Isso estava acontecendo porque os daemons no cluster não conseguiam comunicar os dados do status de integridade de um nó aos componentes internos do Amazon EMR.
- Aprimorados os daemons do EMR no cluster para rastrear corretamente os estados dos nós quando são reutilizados endereços IP para melhorar a confiabilidade durante operações de escalabilidade.
- SPARK-29683. Corrigido um problema em que ocorriam falhas de trabalho durante a redução da escala verticalmente do cluster, pois o Spark presumia que todos os nós disponíveis estavam na lista de negação.
- YARN-9011. Corrigido um problema em que ocorriam falhas de trabalho devido a uma condição de corrida na desativação do YARN quando o cluster tentava aumentar ou reduzir a escala verticalmente.

- Corrigido problema com falhas de etapas ou tarefas durante a escalabilidade do cluster ao ser garantido que os estados dos nós fossem sempre consistentes entre os daemons do Amazon EMR no cluster e o YARN/HDFS.
- Corrigido um problema em que operações de cluster, como redução de escala verticalmente e envio de etapas, falhavam para clusters do Amazon EMR habilitados com a autenticação Kerberos. Isso ocorreu porque o daemon no cluster do Amazon EMR não renovou o tíquete Kerberos, que é necessário para se comunicar com segurança com a execução no nó primário. HDFS/YARN
- As versões mais recentes do Amazon EMR corrigem o problema com um limite menor de “Máximo de arquivos abertos” em relação às versões mais antigas AL2 no Amazon EMR. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR agora incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”.
- O HTTPS agora está habilitado por padrão para repositórios do Amazon Linux. Se você estiver usando uma política de VPCE do Amazon S3 para restringir o acesso a buckets específicos, deverá adicionar o novo ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` do bucket do Amazon Linux à sua política (substitua `$region` pela região em que o endpoint está situado). Para obter mais informações, consulte esse tópico nos fóruns de AWS discussão. Anúncio: o Amazon Linux 2 agora oferece suporte à capacidade de usar HTTPS ao se conectar a repositórios de pacotes .

Problemas conhecidos

- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
  - O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que

existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.

Versões de componente 5.31.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.0	SDK do Amazon SageMaker Spark
emr-ddb	4.15.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.13.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.

Componente	Versão	Descrição
emr-s3-dist-cp	2.15.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.6.0	EMR S3Select Connector
emrfs	2.43.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.11.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.11.0	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.10.0-amzn-0.1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.

Componente	Versão	Descrição
hadoop-hdfs-datanode	2.10.0-amzn-0.1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.10.0-amzn-0.1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.10.0-amzn-0.1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	2.10.0-amzn-0.1	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	2.10.0-amzn-0.1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.10.0-amzn-0.1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.10.0-amzn-0.1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.10.0-amzn-0.1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.10.0-amzn-0.1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.

Componente	Versão	Descrição
hadoop-yarn-timeline-server	2.10.0-amzn-0.1	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.4.13	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.13	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.13	HBase cliente de linha de comando.
hbase-rest-server	1.4.13	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.13	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.7-amzn-1	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.7-amzn-1	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.3.7-amzn-1	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.

Componente	Versão	Descrição
hive-client	2.3.7-amzn-1	O cliente da linha de comando do Hive.
hive-hbase	2.3.7-amzn-1	Cliente do Hive-hbase.
hive-metastore-server	2.3.7-amzn-1	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.7-amzn-1	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.6.0-amzn-0	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-spark	0.6.0-amzn-0	Biblioteca de pacotes para executar o Spark com o Hudi.
hudi-presto	0.6.0-amzn-0	Biblioteca de pacotes para executar o Presto com o Hudi.
hue-server	4.7.1	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	1.1.0	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.0-incubating	Interface REST para interagir com o Apache Spark

Componente	Versão	Descrição
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.6.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.64+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.0	O cliente da linha de comando do Oozie.
oozie-server	5.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	4.3.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.3- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.3- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica

Componente	Versão	Descrição
presto-coordinator	0.238.3-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.238.3-amzn-0	O serviço que executa partes de uma consulta.
presto-client	0.238.3-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.3	O projeto R para computação estatística
ranger-kms-server	1.2.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	2.4.6-amzn-0.1	Os clientes da linha de comando do Spark.
spark-history-server	2.4.6-amzn-0.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.6-amzn-0.1	O mecanismo de execução em memória para o YARN.

Componente	Versão	Descrição
spark-yarn-slave	2.4.6-amzn-0.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.1.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.8.2	O notebook baseado na web que permite a análise de dados interativa.
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.31.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.31.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase

Classificações	Descrição
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive- site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.pro perties do Hive.

Classificações	Descrição
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.

Classificações	Descrição
hudi-env	Altere os valores no ambiente do Hudi.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.

Classificações	Descrição
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.

Classificações	Descrição
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-memory	Alterar os valores no arquivo memory.pr operties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.p roperties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.propert ies do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.

Classificações	Descrição
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.
recordserver-env	Altere os valores no ambiente do EMR. RecordServer
recordserver-conf	Altere os valores no arquivo erver.properties RecordServer do EMR.
recordserver-log4j	Altere os valores no arquivo log4j.properties RecordServer do EMR.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.

Classificações	Descrição
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.31.0 do Amazon EMR

Versões 5.31.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterHub, LivyMXNet, Mahout, Oozie, Phoenix, PigPresto, Spark, Sqoop, TensorFlow, TezZeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.31.0	emr-5.30.2	emr-5.30.1	emr-5.30.0
AWS SDK for Java	1.11.852	1.11.759	1.11.759	1.11.759
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.10.0	1.10.0	1.10.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.7	2.3.6	2.3.6	2.3.6
Hadoop	2.10.0	2.8.5	2.8.5	2.8.5
Hive	2.3.7	2.3.6	2.3.6	2.3.6
Hudi	0.6.0-amzn-0	0.5.2-incubating	0.5.2-incubating	0.5.2-incubating
Hue	4.7.1	4.6.0	4.6.0	4.6.0

	emr-5.31.0	emr-5.30.2	emr-5.30.1	emr-5.30.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.6.0	1.5.1	1.5.1	1.5.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,238,3	0.232	0.232	0.232
Spark	2.4.6	2.4.5	2.4.5	2.4.5
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.1.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas da versão 5.31.0

As notas da versão a seguir incluem informações para a versão 5.31.0 do Amazon EMR. As alterações são referentes à versão 5.30.1.

Data da versão inicial: 9 de outubro de 2020

Data da última atualização: 15 de outubro de 2020

Atualizações

- Atualizado o conector do Amazon Glue para a versão 1.13.0
- SDK do Amazon SageMaker Spark atualizado para a versão 1.4.0
- Atualizado o conector do Amazon Kinesis para a versão 3.5.9
- Atualizado para AWS SDK para Java a versão 1.11.852
- Atualizado o Bigtop-tomcat para a versão 8.5.56
- Atualizado o EMRFS para a versão 2.43.0
- Cliente MetricsAndEventsApiGateway EMR atualizado para a versão 1.4.0
- Atualizado o EMR S3 Dist CP para a versão 2.15.0
- Atualizado o EMR S3 Select para a versão 1.6.0
- Atualizado o Flink para a versão 1.11.0
- Atualizado o Hadoop para a versão 2.10.0
- Atualizado o Hive para a versão 2.3.7
- Atualizado o Hudi para a versão 0.6.0
- Atualizado o Hue para a versão 4.7.1
- Atualizado JupyterHub para a versão 1.1.0
- Atualizado o Mxnet para a versão 1.6.0
- Atualizado o OpenCV para a versão 4.3.0
- Atualizado o Presto para a versão 0.238.3
- Atualizado TensorFlow para a versão 2.1.0

Alterações, melhorias e problemas resolvidos

- Esta é uma versão para corrigir problemas com o Amazon EMR Scaling quando ele não consegue up/scale reduzir a escala de um cluster com sucesso ou causa falhas no aplicativo.

- Corrigido um problema em que as solicitações de escalabilidade falhavam em um cluster grande e altamente utilizado quando os daemons do Amazon EMR no cluster estavam executando atividades de verificação de integridade, como a coleta do estado do nó do YARN e o estado do nó do HDFS. Isso estava acontecendo porque os daemons no cluster não conseguiam comunicar os dados do status de integridade de um nó aos componentes internos do Amazon EMR.
- Aprimorados os daemons do EMR no cluster para rastrear corretamente os estados dos nós quando são reutilizados endereços IP para melhorar a confiabilidade durante operações de escalabilidade.
- SPARK-29683. Corrigido um problema em que ocorriam falhas de trabalho durante a redução da escala verticalmente do cluster, pois o Spark presumia que todos os nós disponíveis estavam na lista de negação.
- YARN-9011. Corrigido um problema em que ocorriam falhas de trabalho devido a uma condição de corrida na desativação do YARN quando o cluster tentava aumentar ou reduzir a escala verticalmente.
- Corrigido problema com falhas de etapas ou tarefas durante a escalabilidade do cluster ao ser garantido que os estados dos nós fossem sempre consistentes entre os daemons do Amazon EMR no cluster e o YARN/HDFS.
- Corrigido um problema em que operações de cluster, como redução de escala verticalmente e envio de etapas, falhavam para clusters do Amazon EMR habilitados com a autenticação Kerberos. Isso ocorreu porque o daemon no cluster do Amazon EMR não renovou o tíquete Kerberos, que é necessário para se comunicar com segurança com a execução no nó primário. HDFS/YARN
- As versões mais recentes do Amazon EMR corrigem o problema com um limite menor de “Máximo de arquivos abertos” em relação às versões mais antigas AL2 no Amazon EMR. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR agora incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”.
- As estatísticas de colunas do Hive são compatíveis com as versões 5.31.0 e posteriores do Amazon EMR.
- Atualizadas versões do componente.
- Suporte ao EMRFS S3EC V2 no Amazon EMR 5.31.0. Nas versões 1.11.837 e posteriores do SDK para Java no S3, a versão 2 do cliente de criptografia (S3EC V2) foi introduzida com vários aprimoramentos de segurança. Para obter mais informações, consulte:
  - Publicações no blog do S3: Atualizações no cliente de criptografia do Amazon S3.

- AWS SDK para Java Guia do desenvolvedor: Migre clientes de criptografia e descriptografia para a V2.
- Guia de gerenciamento do EMR: Criptografia do lado do cliente do Amazon S3.

O Encryption Client V1 ainda está disponível no SDK para compatibilidade retroativa.

Novos recursos

- Limite inferior de “Máximo de arquivos abertos” em versões mais antigas AL2 [corrigido em versões mais recentes]. Versões do Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 e emr-6.2.0 são baseadas em versões mais antigas do Amazon Linux 2 (), que AL2 têm uma configuração de limite inferior para “Máximo de arquivos abertos” quando clusters do Amazon EMR são criados com a AMI padrão. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”. Versões com o limite inferior de arquivos abertos causam o erro “Muitos arquivos abertos” ao ser enviado um trabalho do Spark. Nas versões afetadas, a AMI padrão do Amazon EMR tem uma configuração de ulimit padrão de 4096 para “Máximo de arquivos abertos”, que é inferior ao limite de 65536 arquivos na AMI mais recente do Amazon Linux 2. A configuração inferior de ulimit para “Máximo de arquivos abertos” causa falhas em trabalhos do Spark quando o driver e o executor do Spark tentam abrir mais de 4096 arquivos. Para corrigir o problema, o Amazon EMR tem um script de ação de bootstrap (BA) que ajusta a configuração de ulimit na criação do cluster.

  Se você está usando uma versão mais antiga do Amazon EMR que não tem a correção permanente para esse problema, a solução alternativa a seguir permite que você defina explicitamente o ulimit instance-controller para um máximo de 65536 arquivos.

  Defina explicitamente um ulimit na linha de comando

  1. Edite `/etc/systemd/system/instance-controller.service` para adicionar os seguintes parâmetros à seção Serviço.

     ```
     LimitNOFILE=65536

     LimitNPROC=65536
     ```

  2. Reiniciar InstanceController

     ```
     $ sudo systemctl daemon-reload
     ```

```
$ sudo systemctl restart instance-controller
```

Defina um ulimit usando a ação de bootstrap (BA)

Você também pode usar um script de ação de bootstrap (BA) para configurar o ulimit instance-controller para 65536 arquivos na criação do cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- Com a versão 5.31.0 do Amazon EMR, você pode iniciar um cluster que se integre ao Lake Formation. Essa integração fornece filtragem de dados refinada em nível de coluna para bancos de dados e tabelas no Glue Data Catalog. AWS Ela também permite logon único federado para Cadernos do EMR ou para o Apache Zeppelin em um sistema de identidade empresarial. Para obter mais informações, consulte Integrating Amazon EMR with AWS Lake Formation no Guia de gerenciamento do Amazon EMR.

  Atualmente, o Amazon EMR com Lake Formation está disponível em 16 AWS regiões: Leste dos EUA (Ohio e Norte da Virgínia), Oeste dos EUA (Norte da Califórnia e Oregon), Ásia-Pacífico (Mumbai, Seul, Cingapura, Sydney e Tóquio), Canadá (Central), Europa (Frankfurt, Irlanda, Londres, Paris e Estocolmo) e América do Sul (São Paulo).

Problemas conhecidos

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

  Solução:

  - SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
  - Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

- Quando AtRestEncryption a criptografia HDFS é habilitada em um cluster que usa o Amazon EMR 5.31.0 ou 5.32.0, as consultas do Hive resultam na seguinte exceção de tempo de execução.

  ```
  TaskAttempt 3 failed, info=[Error: Error while running task ( failure ) :
   attempt_1604112648850_0001_1_01_000000_3:java.lang.RuntimeException:
   java.lang.RuntimeException: Hive Runtime Error while closing
   operators: java.io.IOException: java.util.ServiceConfigurationError:
   org.apache.hadoop.security.token.TokenIdentifier: Provider
   org.apache.hadoop.hbase.security.token.AuthenticationTokenIdentifier not found
  ```

- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
  - O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

  Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.

Versões de componente 5.31.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.4.0	SDK do Amazon SageMaker Spark
emr-ddb	4.15.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.13.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.15.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.6.0	EMR S3Select Connector
emrfs	2.43.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.11.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
flink-jobmanager-config	1.11.0	Gerenciando recursos nos nós do EMR para o Apache Flink. JobManager
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.10.0-amzn-0	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.10.0-amzn-0	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.10.0-amzn-0	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.10.0-amzn-0	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	2.10.0-amzn-0	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	2.10.0-amzn-0	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.10.0-amzn-0	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider

Componente	Versão	Descrição
hadoop-mapred	2.10.0-amzn-0	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.10.0-amzn-0	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.10.0-amzn-0	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hadoop-yarn-timeline-server	2.10.0-amzn-0	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.4.13	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.13	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.13	HBase cliente de linha de comando.
hbase-rest-server	1.4.13	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.13	Serviço que fornece um endpoint Thrift para. HBase

Componente	Versão	Descrição
hcatalog-client	2.3.7-amzn-1	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.7-amzn-1	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.3.7-amzn-1	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.7-amzn-1	O cliente da linha de comando do Hive.
hive-hbase	2.3.7-amzn-1	Cliente do Hive-hbase.
hive-metastore-server	2.3.7-amzn-1	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.7-amzn-1	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.6.0-amzn-0	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-spark	0.6.0-amzn-0	Biblioteca de pacotes para executar o Spark com o Hudi.

Componente	Versão	Descrição
hudi-presto	0.6.0-amzn-0	Biblioteca de pacotes para executar o Presto com o Hudi.
hue-server	4.7.1	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	1.1.0	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.6.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.64	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.0	O cliente da linha de comando do Oozie.
oozie-server	5.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.

Componente	Versão	Descrição
opencv	4.3.0	Biblioteca de Visão Computacional de Código Aberto.
phoenix-library	4.14.3- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.3- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.238.3-amzn-0	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.238.3-amzn-0	O serviço que executa partes de uma consulta.
presto-client	0.238.3-amzn-0	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.3	O projeto R para computação estatística
ranger-kms-server	1.2.0	Sistema de gerenciamento de chaves Apache Ranger

Componente	Versão	Descrição
spark-client	2.4.6-amzn-0	Os clientes da linha de comando do Spark.
spark-history-server	2.4.6-amzn-0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.6-amzn-0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.4.6-amzn-0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	2.1.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.8.2	O notebook baseado na web que permite a análise de dados interativa.

Componente	Versão	Descrição
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.31.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.31.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.

Classificações	Descrição
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.

Classificações	Descrição
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.

Classificações	Descrição
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
hudi-env	Altere os valores no ambiente do Hudi.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.

Classificações	Descrição
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.

Classificações	Descrição
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.

Classificações	Descrição
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.
recordserver-env	Altere os valores no ambiente do EMR. RecordServer

Classificações	Descrição
recordserver-conf	Altere os valores no arquivo erver.properties RecordServer do EMR.
recordserver-log4j	Altere os valores no arquivo log4j.properties RecordServer do EMR.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.con f do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.

Classificações	Descrição
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.30.2 do Amazon EMR

Versões 5.30.2 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterHub, LivyMXNet, Mahout, Oozie, Phoenix, PigPresto, Spark, Sqoop, TensorFlow, TezZeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.30.2	emr-5.30.1	emr-5.30.0	emr-5.29.0
AWS SDK for Java	1.11.759	1.11.759	1.11.759	1.11.682
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.6

	emr-5.30.2	emr-5.30.1	emr-5.30.0	emr-5.29.0
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.10.0	1.10.0	1.10.0	1.9.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.10
HCatalog	2.3.6	2.3.6	2.3.6	2.3.6
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.6	2.3.6	2.3.6
Hudi	0.5.2-incubating	0.5.2-incubating	0.5.2-incubating	0.5.0-incubating
Hue	4.6.0	4.6.0	4.6.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.0.0
Livy	0.7.0	0.7.0	0.7.0	0.6.0
MXNet	1.5.1	1.5.1	1.5.1	1.5.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.1.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3

	emr-5.30.2	emr-5.30.1	emr-5.30.0	emr-5.29.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.232	0.232	0.232	0.227
Spark	2.4.5	2.4.5	2.4.5	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas da versão 5.30.2

Esta é uma versão para corrigir problemas com o Amazon EMR Scaling quando ele não consegue up/scale reduzir a escala de um cluster com sucesso ou causa falhas no aplicativo.

Alterações, melhorias e problemas resolvidos

- Corrigido um problema em que as solicitações de escalabilidade falhavam em um cluster grande e altamente utilizado quando os daemons do Amazon EMR no cluster estavam executando atividades de verificação de integridade, como a coleta do estado do nó do YARN e o estado do nó do HDFS. Isso estava acontecendo porque os daemons no cluster não conseguiam comunicar os dados do status de integridade de um nó aos componentes internos do Amazon EMR.
- Aprimorados os daemons do EMR no cluster para rastrear corretamente os estados dos nós quando são reutilizados endereços IP para melhorar a confiabilidade durante operações de escalabilidade.
- SPARK-29683. Corrigido um problema em que ocorriam falhas de trabalho durante a redução da escala verticalmente do cluster, pois o Spark presumia que todos os nós disponíveis estavam na lista de negação.

- YARN-9011. Corrigido um problema em que ocorriam falhas de trabalho devido a uma condição de corrida na desativação do YARN quando o cluster tentava aumentar ou reduzir a escala verticalmente.
- Corrigido problema com falhas de etapas ou tarefas durante a escalabilidade do cluster ao ser garantido que os estados dos nós fossem sempre consistentes entre os daemons do Amazon EMR no cluster e o YARN/HDFS.
- Corrigido um problema em que operações de cluster, como redução de escala verticalmente e envio de etapas, falhavam para clusters do Amazon EMR habilitados com a autenticação Kerberos. Isso ocorreu porque o daemon no cluster do Amazon EMR não renovou o tíquete Kerberos, que é necessário para se comunicar com segurança com a execução no nó primário. HDFS/YARN
- As versões mais recentes do Amazon EMR corrigem o problema com um limite menor de “Máximo de arquivos abertos” em relação às versões mais antigas AL2 no Amazon EMR. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR agora incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”.
- O HTTPS agora está habilitado por padrão para repositórios do Amazon Linux. Se você estiver usando uma política de VPCE do Amazon S3 para restringir o acesso a buckets específicos, deverá adicionar o novo ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` do bucket do Amazon Linux à sua política (substitua `$region` pela região em que o endpoint está situado). Para obter mais informações, consulte esse tópico nos fóruns de AWS discussão. Anúncio: o Amazon Linux 2 agora oferece suporte à capacidade de usar HTTPS ao se conectar a repositórios de pacotes .

Problemas conhecidos

- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.

- O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere / (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.

Versões de componente 5.30.2

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.3.0	SDK do Amazon SageMaker Spark
emr-ddb	4.14.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.13.0	Bibliotecas convenientes para o ecossistema do Hadoop.

Componente	Versão	Descrição
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.14.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.5.0	EMR S3Select Connector
emrfs	2.40.0	O conector do Amazon S3 para aplicativos do ecossiste ma do Hadoop.
flink-client	1.10.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.5-amzn-6.1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.

Componente	Versão	Descrição
hadoop-hdfs-datanode	2.8.5-amzn-6.1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.5-amzn-6.1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.5-amzn-6.1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	2.8.5-amzn-6.1	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	2.8.5-amzn-6.1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.5-amzn-6.1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.8.5-amzn-6.1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.5-amzn-6.1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-6.1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.

Componente	Versão	Descrição
hadoop-yarn-timeline-server	2.8.5-amzn-6.1	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.4.13	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.13	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.13	HBase cliente de linha de comando.
hbase-rest-server	1.4.13	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.13	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.6-amzn-2	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.6-amzn-2	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.3.6-amzn-2	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.

Componente	Versão	Descrição
hive-client	2.3.6-amzn-2	O cliente da linha de comando do Hive.
hive-hbase	2.3.6-amzn-2	Cliente do Hive-hbase.
hive-metastore-server	2.3.6-amzn-2	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.6-amzn-2	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.5.2-incubating	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.5.2-incubating	Biblioteca de pacotes para executar o Presto com o Hudi.
hue-server	4.6.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	1.1.0	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP

Componente	Versão	Descrição
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.5.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.64+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.0	O cliente da linha de comando do Oozie.
oozie-server	5.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.3- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.3- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.232	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.

Componente	Versão	Descrição
presto-worker	0.232	O serviço que executa partes de uma consulta.
presto-client	0.232	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.3	O projeto R para computação estatística
ranger-kms-server	1.2.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	2.4.5-amzn-0.1	Os clientes da linha de comando do Spark.
spark-history-server	2.4.5-amzn-0.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.5-amzn-0.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.4.5-amzn-0.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.

Componente	Versão	Descrição
tensorflow	1.14.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.8.2	O notebook baseado na web que permite a análise de dados interativa.
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.30.2

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.30.2

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.

Classificações	Descrição
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.

Classificações	Descrição
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.

Classificações	Descrição
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.proper ties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.x ml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properti es do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
hudi-env	Altere os valores no ambiente do Hudi.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_n otebook_config.py do Notebook Jupyter.

Classificações	Descrição
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.

Classificações	Descrição
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.

Classificações	Descrição
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.

Classificações	Descrição
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.
recordserver-env	Altere os valores no ambiente do EMR. RecordServer
recordserver-conf	Altere os valores no arquivo erver.properties RecordServer do EMR.
recordserver-log4j	Altere os valores no arquivo log4j.properties RecordServer do EMR.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.

Classificações	Descrição
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.30.1 do Amazon EMR

Versões 5.30.1 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterHub, LivyMXNet, Mahout, Oozie, Phoenix, PigPresto, Spark, Sqoop, TensorFlow, TezZeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.30.1	emr-5.30.0	emr-5.29.0	emr-5.28.1
AWS SDK for Java	1.11.759	1.11.759	1.11.682	1.11.659
Python	2.7, 3.7	2.7, 3.7	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.10.0	1.10.0	1.9.1	1.9.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.10	1.4.10
HCatalog	2.3.6	2.3.6	2.3.6	2.3.6
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.6	2.3.6	2.3.6
Hudi	0.5.2-incubating	0.5.2-incubating	0.5.0-incubating	0.5.0-incubating
Hue	4.6.0	4.6.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.0.0	1.0.0
Livy	0.7.0	0.7.0	0.6.0	0.6.0
MXNet	1.5.1	1.5.1	1.5.1	1.5.1

	emr-5.30.1	emr-5.30.0	emr-5.29.0	emr-5.28.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.1.0	5.1.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.232	0.232	0.227	0.227
Spark	2.4.5	2.4.5	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas da versão 5.30.1

As notas da versão a seguir incluem informações para a versão 5.30.1 do Amazon EMR. As alterações são referentes à versão 5.30.0.

Data da versão inicial: 30 de junho de 2020

Data da última atualização: 24 de agosto de 2020

Alterações, melhorias e problemas resolvidos

- As versões mais recentes do Amazon EMR corrigem o problema com um limite menor de “Máximo de arquivos abertos” em relação às versões mais antigas AL2 no Amazon EMR. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR

agora incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”.

- Corrigido um problema em que o processo do controlador da instância gerava um número infinito de processos.
- Corrigido um problema em que o Hue não conseguia executar uma consulta do Hive, mostrando a mensagem “o banco de dados está bloqueado” e impedindo a execução de consultas.
- Corrigido um problema no Spark para permitir que mais tarefas fossem executadas simultaneamente no cluster do EMR.
- Corrigido um problema no caderno Jupyter que causava um “erro de muitos arquivos abertos” no servidor Jupyter.
- Corrigido um problema com as horas de início do cluster.

Novos recursos

- A interface do usuário do Tez e as interfaces de aplicações persistentes do servidor de linha do tempo do YARN estão disponíveis com as versões 6.x do Amazon EMR e com as versões 5.30.1 e posteriores do EMR. O acesso com um clique no link do histórico persistente da aplicação permite que você acesse rapidamente o histórico de tarefas sem configurar um proxy da Web por meio de uma conexão SSH. Os logs de clusters ativos e encerrados ficam disponíveis por 30 dias após o término da aplicação. Para obter mais informações, consulte View Persistent Application User Interfaces no Guia de gerenciamento do Amazon EMR.
- A execução do EMR Notebook APIs está disponível para executar notebooks EMR por meio de um script ou linha de comando. A capacidade de iniciar, parar, listar e descrever as execuções do notebook EMR sem o AWS console permite que você controle programaticamente um notebook EMR. Ao usar uma célula do caderno parametrizada, você pode passar valores de parâmetros diferentes para um caderno sem precisar criar uma cópia do caderno para cada novo conjunto de valores de parâmetros. Consulte Ações de API do EMR. Para obter um exemplo de código, consulte Exemplos de comandos para executar Cadernos do EMR programaticamente.

Problemas conhecidos

- Limite inferior de “Máximo de arquivos abertos” em versões mais antigas AL2 [corrigido em versões mais recentes]. Versões do Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 e emr-6.2.0 são baseadas em versões mais antigas do Amazon Linux 2 (), que AL2 têm uma configuração de limite inferior para “Máximo de arquivos abertos” quando clusters do

Amazon EMR são criados com a AMI padrão. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”. Versões com o limite inferior de arquivos abertos causam o erro “Muitos arquivos abertos” ao ser enviado um trabalho do Spark. Nas versões afetadas, a AMI padrão do Amazon EMR tem uma configuração de ulimit padrão de 4096 para “Máximo de arquivos abertos”, que é inferior ao limite de 65536 arquivos na AMI mais recente do Amazon Linux 2. A configuração inferior de ulimit para “Máximo de arquivos abertos” causa falhas em trabalhos do Spark quando o driver e o executor do Spark tentam abrir mais de 4096 arquivos. Para corrigir o problema, o Amazon EMR tem um script de ação de bootstrap (BA) que ajusta a configuração de ulimit na criação do cluster.

Se você está usando uma versão mais antiga do Amazon EMR que não tem a correção permanente para esse problema, a solução alternativa a seguir permite que você defina explicitamente o ulimit instance-controller para um máximo de 65536 arquivos.

Defina explicitamente um ulimit na linha de comando

1. Edite `/etc/systemd/system/instance-controller.service` para adicionar os seguintes parâmetros à seção Serviço.

   ```
   LimitNOFILE=65536

   LimitNPROC=65536
   ```

2. Reiniciar InstanceController

   ```
   $ sudo systemctl daemon-reload

   $ sudo systemctl restart instance-controller
   ```

Defina um ulimit usando a ação de bootstrap (BA)

Você também pode usar um script de ação de bootstrap (BA) para configurar o ulimit instance-controller para 65536 arquivos na criação do cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
```

```
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- Cadernos do EMR

  O atributo que permite instalar kernels e bibliotecas Python adicionais no nó primário do cluster está desabilitado por padrão na versão 5.30.1 do EMR. Para obter mais informações sobre esse atributo, consulte Instalar kernels e bibliotecas Python em um nó primário do cluster.

  Para habilitar o recurso, faça o seguinte:

  1. Certifique-se de que a política de permissões anexada ao perfil de serviço para os Cadernos do EMR permite a seguinte ação:

     ```
     elasticmapreduce:ListSteps
     ```

     Para obter mais informações, consulte Função de serviço do EMR Notebooks.

  2. Use o AWS CLI para executar uma etapa no cluster que configura os Notebooks EMR, conforme mostrado no exemplo a seguir. *us-east-1*Substitua pela região em que seu cluster reside. Para obter mais informações, consulte Adding Steps to a Cluster Using the AWS CLI.

     ```
     aws emr add-steps  --cluster-id MyClusterID --steps
      Type=CUSTOM_JAR,Name=EMRNotebooksSetup,ActionOnFailure=CONTINUE,Jar=s3://us-
     east-1.elasticmapreduce/libs/script-runner/script-runner.jar,Args=["s3://
     awssupportdatasvcs.com/bootstrap-actions/EMRNotebooksSetup/emr-notebooks-setup.sh"]
     ```

- Ajuste de escala gerenciado

  As operações de ajuste de escala gerenciado nos clusters das versões 5.30.0 e 5.30.1 sem o Presto instalado podem causar falhas na aplicação ou fazer com que um grupo de instâncias ou uma frota de instâncias uniforme permaneça no estado ARRESTED, sobretudo quando uma

operação de redução da escala verticalmente logo é seguida por uma operação de aumento da escala verticalmente.

Como solução alternativa, escolha o Presto como uma aplicação a ser instalada ao criar um cluster com as versões 5.30.0 e 5.30.1 do Amazon EMR, mesmo que o trabalho não exija o Presto.

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

  Solução:

  - SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
  - Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.

- Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
- O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.

Versões de componente 5.30.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.3.0	SDK do Amazon SageMaker Spark
emr-ddb	4.14.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.

Componente	Versão	Descrição
emr-goodies	2.13.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.14.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.5.0	EMR S3Select Connector
emrfs	2.40.0	O conector do Amazon S3 para aplicativos do ecossiste ma do Hadoop.
flink-client	1.10.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.

Componente	Versão	Descrição
hadoop-client	2.8.5-amzn-6	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.5-amzn-6	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.5-amzn-6	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.5-amzn-6	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	2.8.5-amzn-6	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	2.8.5-amzn-6	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.5-amzn-6	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-6	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.5-amzn-6	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.8.5-amzn-6	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.5-amzn-6	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.4.13	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.13	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.13	HBase cliente de linha de comando.
hbase-rest-server	1.4.13	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.13	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.6-amzn-2	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.6-amzn-2	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	2.3.6-amzn-2	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.6-amzn-2	O cliente da linha de comando do Hive.
hive-hbase	2.3.6-amzn-2	Cliente do Hive-hbase.
hive-metastore-server	2.3.6-amzn-2	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.6-amzn-2	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.5.2-incubating	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.5.2-incubating	Biblioteca de pacotes para executar o Presto com o Hudi.
hue-server	4.6.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	1.1.0	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.0-incubating	Interface REST para interagir com o Apache Spark

Componente	Versão	Descrição
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.5.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.64	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.0	O cliente da linha de comando do Oozie.
oozie-server	5.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.3- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.3- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica

Componente	Versão	Descrição
presto-coordinator	0.232	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.232	O serviço que executa partes de uma consulta.
presto-client	0.232	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.3	O projeto R para computação estatística
ranger-kms-server	1.2.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	2.4.5-amzn-0	Os clientes da linha de comando do Spark.
spark-history-server	2.4.5-amzn-0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.5-amzn-0	O mecanismo de execução em memória para o YARN.

Componente	Versão	Descrição
spark-yarn-slave	2.4.5-amzn-0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	1.14.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.8.2	O notebook baseado na web que permite a análise de dados interativa.
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.30.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.30.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase

Classificações	Descrição
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive- site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.pro perties do Hive.

Classificações	Descrição
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.

Classificações	Descrição
hudi-env	Altere os valores no ambiente do Hudi.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.

Classificações	Descrição
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.

Classificações	Descrição
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-memory	Alterar os valores no arquivo memory.pr operties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.p roperties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.propert ies do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.

Classificações	Descrição
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.
recordserver-env	Altere os valores no ambiente do EMR. RecordServer
recordserver-conf	Altere os valores no arquivo erver.properties RecordServer do EMR.
recordserver-log4j	Altere os valores no arquivo log4j.properties RecordServer do EMR.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.

Classificações	Descrição
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.30.0 do Amazon EMR

Versões 5.30.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterHub, LivyMXNet, Mahout, Oozie, Phoenix, PigPresto, Spark, Sqoop, TensorFlow, TezZeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.30.0	emr-5.29.0	emr-5.28.1	emr-5.28.0
AWS SDK for Java	1.11.759	1.11.682	1.11.659	1.11.659
Python	2.7, 3.7	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.10.0	1.9.1	1.9.0	1.9.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.10	1.4.10	1.4.10
HCatalog	2.3.6	2.3.6	2.3.6	2.3.6
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.6	2.3.6	2.3.6
Hudi	0.5.2-incubating	0.5.0-incubating	0.5.0-incubating	0.5.0-incubating
Hue	4.6.0	4.4.0	4.4.0	4.4.0

	emr-5.30.0	emr-5.29.0	emr-5.28.1	emr-5.28.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.0.0	1.0.0	1.0.0
Livy	0.7.0	0.6.0	0.6.0	0.6.0
MXNet	1.5.1	1.5.1	1.5.1	1.5.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.232	0.227	0.227	0.227
Spark	2.4.5	2.4.4	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas da versão 5.30.0

As notas da versão a seguir incluem informações para a versão 5.30.0 do Amazon EMR. As alterações são referentes à versão 5.29.0.

Data da versão inicial: 13 de maio de 2020

Data da última atualização: 25 de junho de 2020

Atualizações

- Atualizado para AWS SDK para Java a versão 1.11.759
- SDK do Amazon SageMaker Spark atualizado para a versão 1.3.0
- Atualização do EMR Record Server para a versão 1.6.0
- Atualização do Flink para a versão 1.10.0
- Atualização do Ganglia para a versão 3.7.2
- Atualizado HBase para a versão 1.4.13
- Atualização do Hudi para a versão 0.5.2-incubating
- Atualização do Hue para a versão 4.6.0
- Atualizado JupyterHub para a versão 1.1.0
- Atualização do Livy para a versão 0.7.0-incubating
- Atualização do Oozie para a versão 5.2.0
- Atualização do Presto para a versão 0.232
- Atualização do Spark para a versão 2.4.5
- Atualização de conectores e drivers: Amazon Glue Connector 1.12.0; Amazon Kinesis Connector 3.5.0; EMR DynamoDB Connector 4.14.0

Novos recursos

- Cadernos do EMR: quando usados com clusters do EMR criados com a versão 5.30.0, os kernels dos Cadernos do EMR são executados no cluster. Isso melhora o desempenho do bloco de anotações e permite que instalar e personalizar kernels. Você também pode instalar bibliotecas Python no nó primário do cluster. Para obter mais informações, consulte Instalar e usar kernels e bibliotecas no EMR Management Guide.
- Ajuste de escala gerenciado: com as versões 5.30.0 e posteriores do Amazon EMR, é possível habilitar o ajuste de escala gerenciado pelo EMR para aumentar ou diminuir automaticamente

o número de instâncias ou unidades no cluster com base na workload. O Amazon EMR avalia continuamente as métricas do cluster para tomar decisões de ajuste de escala que otimizam os clusters em termos de custo e velocidade. Para obter mais informações, consulte Escalar recursos de cluster no Guia de gerenciamento do Amazon EMR.

- Criptografe arquivos de log armazenados no Amazon S3 — Com o Amazon EMR versão 5.30.0 e posterior, você pode criptografar arquivos de log armazenados no Amazon S3 com uma chave gerenciada pelo cliente. AWS KMS Para obter mais informações, consulte Encrypt log files stored in Amazon S3 no Guia de gerenciamento do Amazon EMR.
- Suporte ao Amazon Linux 2: nas versões 5.30.0 e posteriores do EMR, o EMR usa o sistema operacional Amazon Linux 2. A nova personalização AMIs (Amazon Machine Image) deve ser baseada na Amazon Linux 2 AMI. Para obter mais informações, consulte Usando uma AMI personalizada.
- Escalabilidade automática tranquila do Presto: os clusters do EMR que usam a versão 5.30.0 podem ser definidos com um período limite de ajuste de escala automático que dá tempo às tarefas do Presto para concluir a execução antes do nó delas ser desativado. Para obter mais informações, consulte Usar a escalabilidade automática do Presto com desativação tranquila.
- Criação de instância de frota com nova opção de estratégia de alocação: uma nova opção de estratégia de alocação está disponível nas versões 5.12.1 e posteriores do EMR. Ele oferece provisionamento de cluster mais rápido, alocação de spot mais precisa e menos interrupção de instâncias spot. São necessárias atualizações para perfis de serviço do EMR não padrão. Consulte Configurar frotas de instâncias.
- Comandos sudo systemctl stop e sudo systemctl start: nas versões 5.30.0 e posteriores do EMR, que usam o sistema operacional Amazon Linux 2, o EMR usa os comandos `sudo systemctl stop` e `sudo systemctl start` para reiniciar serviços. Para obter mais informações, consulte Como reinicio um serviço no Amazon EMR?.

Alterações, melhorias e problemas resolvidos

- O EMR versão 5.30.0 não instala o Ganglia por padrão. É possível selecionar explicitamente o Ganglia para ser instalado ao criar um cluster.
- Otimizações do desempenho do Spark
- Otimizações do desempenho do Presto
- O Python 3 é o padrão para as versões 5.30.0 e posteriores do Amazon EMR.
- O grupo de segurança gerenciado padrão para acesso ao serviço em sub-redes privadas foi atualizado com novas regras. Se você usar um grupo de segurança personalizado para acesso

ao serviço, será necessário incluir as mesmas regras do grupo de segurança gerenciado padrão. Para obter mais informações, consulte Grupo de segurança gerenciado pelo Amazon EMR para acesso ao serviço (sub-redes privadas). Se você usar um perfil de serviço personalizado para o Amazon EMR, será necessário conceder permissão para `ec2:describeSecurityGroups` de modo que o EMR possa validar se os grupos de segurança são criados corretamente. Se você usar o `EMR_DefaultRole`, essa permissão já estará incluída na política gerenciada padrão.

Problemas conhecidos

- Limite inferior de “Máximo de arquivos abertos” em versões mais antigas AL2 [corrigido em versões mais recentes]. Versões do Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 e emr-6.2.0 são baseadas em versões mais antigas do Amazon Linux 2 (), que AL2 têm uma configuração de limite inferior para “Máximo de arquivos abertos” quando clusters do Amazon EMR são criados com a AMI padrão. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”. Versões com o limite inferior de arquivos abertos causam o erro “Muitos arquivos abertos” ao ser enviado um trabalho do Spark. Nas versões afetadas, a AMI padrão do Amazon EMR tem uma configuração de ulimit padrão de 4096 para “Máximo de arquivos abertos”, que é inferior ao limite de 65536 arquivos na AMI mais recente do Amazon Linux 2. A configuração inferior de ulimit para “Máximo de arquivos abertos” causa falhas em trabalhos do Spark quando o driver e o executor do Spark tentam abrir mais de 4096 arquivos. Para corrigir o problema, o Amazon EMR tem um script de ação de bootstrap (BA) que ajusta a configuração de ulimit na criação do cluster.

  Se você está usando uma versão mais antiga do Amazon EMR que não tem a correção permanente para esse problema, a solução alternativa a seguir permite que você defina explicitamente o ulimit instance-controller para um máximo de 65536 arquivos.

  Defina explicitamente um ulimit na linha de comando

  1. Edite `/etc/systemd/system/instance-controller.service` para adicionar os seguintes parâmetros à seção Serviço.

     ```
     LimitNOFILE=65536

     LimitNPROC=65536
     ```

  2. Reiniciar InstanceController

```
$ sudo systemctl daemon-reload

$ sudo systemctl restart instance-controller
```

Defina um ulimit usando a ação de bootstrap (BA)

Você também pode usar um script de ação de bootstrap (BA) para configurar o ulimit instance-controller para 65536 arquivos na criação do cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- Ajuste de escala gerenciado

  As operações de ajuste de escala gerenciado nos clusters das versões 5.30.0 e 5.30.1 sem o Presto instalado podem causar falhas na aplicação ou fazer com que um grupo de instâncias ou uma frota de instâncias uniforme permaneça no estado `ARRESTED`, sobretudo quando uma operação de redução da escala verticalmente logo é seguida por uma operação de aumento da escala verticalmente.

  Como solução alternativa, escolha o Presto como uma aplicação a ser instalada ao criar um cluster com as versões 5.30.0 e 5.30.1 do Amazon EMR, mesmo que o trabalho não exija o Presto.

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

Solução:

- SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
- Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

- O mecanismo de banco de dados padrão para o Hue 4.6.0 é SQLite, o que causa problemas quando você tenta usar o Hue com um banco de dados externo. Para corrigir isso, defina `engine` na sua classificação de configuração `hue-ini` como `mysql`. Esse problema foi corrigido na versão 5.30.1 do Amazon EMR.
- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.

- Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
- O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.

Versões de componente 5.30.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.3.0	SDK do Amazon SageMaker Spark
emr-ddb	4.14.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.

Componente	Versão	Descrição
emr-goodies	2.13.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.5.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-notebook-env	1.0.0	Conda env para bloco de anotações do EMR
emr-s3-dist-cp	2.14.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.5.0	EMR S3Select Connector
emrfs	2.40.0	O conector do Amazon S3 para aplicativos do ecossiste ma do Hadoop.
flink-client	1.10.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.5-amzn-6	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.5-amzn-6	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.5-amzn-6	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.5-amzn-6	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	2.8.5-amzn-6	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	2.8.5-amzn-6	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.5-amzn-6	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.8.5-amzn-6	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.

Componente	Versão	Descrição
hadoop-yarn-nodemanager	2.8.5-amzn-6	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-6	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hadoop-yarn-timeline-server	2.8.5-amzn-6	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.4.13	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.13	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.13	HBase cliente de linha de comando.
hbase-rest-server	1.4.13	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.13	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.6-amzn-2	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.

Componente	Versão	Descrição
hcatalog-server	2.3.6-amzn-2	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.3.6-amzn-2	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.6-amzn-2	O cliente da linha de comando do Hive.
hive-hbase	2.3.6-amzn-2	Cliente do Hive-hbase.
hive-metastore-server	2.3.6-amzn-2	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.6-amzn-2	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.5.2-incubating	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.5.2-incubating	Biblioteca de pacotes para executar o Presto com o Hudi.
hue-server	4.6.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop

Componente	Versão	Descrição
jupyterhub	1.1.0	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.7.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.5.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mariadb-server	5.5.64	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.2.0	O cliente da linha de comando do Oozie.
oozie-server	5.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.3- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente

Componente	Versão	Descrição
phoenix-query-server	4.14.3- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.232	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.232	O serviço que executa partes de uma consulta.
presto-client	0.232	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.3	O projeto R para computação estatística
ranger-kms-server	1.2.0	Sistema de gerenciamento de chaves Apache Ranger
spark-client	2.4.5-amzn-0	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	2.4.5-amzn-0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.5-amzn-0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.4.5-amzn-0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	1.14.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.8.2	O notebook baseado na web que permite a análise de dados interativa.
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.

Componente	Versão	Descrição
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.30.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.30.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.

Classificações	Descrição
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.

Classificações	Descrição
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.

Classificações	Descrição
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
hudi-env	Altere os valores no ambiente do Hudi.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.

Classificações	Descrição
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.

Classificações	Descrição
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.

Classificações	Descrição
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.
recordserver-env	Altere os valores no ambiente do EMR. RecordServer
recordserver-conf	Altere os valores no arquivo erver.properties RecordServer do EMR.
recordserver-log4j	Altere os valores no arquivo log4j.properties RecordServer do EMR.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.

Classificações	Descrição
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.29.0 do Amazon EMR

Versões 5.29.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterHub, LivyMXNet, Mahout, Oozie, Phoenix, PigPresto, Spark, Sqoop, TensorFlow, TezZeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.29.0	emr-5.28.1	emr-5.28.0	emr-5.27.1
AWS SDK for Java	1.11.682	1.11.659	1.11.659	1.11.615
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.9.1	1.9.0	1.9.0	1.8.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.10	1.4.10	1.4.10

	emr-5.29.0	emr-5.28.1	emr-5.28.0	emr-5.27.1
HCatalog	2.3.6	2.3.6	2.3.6	2.3.5
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.6	2.3.6	2.3.5
Hudi	0.5.0-incubating	0.5.0-incubating	0.5.0-incubating	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.0.0	1.0.0	1.0.0	1.0.0
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.5.1	1.5.1	1.5.1	1.4.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.227	0.227	0.227	0.224
Spark	2.4.4	2.4.4	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-5.29.0	emr-5.28.1	emr-5.28.0	emr-5.27.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas da versão 5.29.0

As notas da versão a seguir incluem informações para a versão 5.29.0 do Amazon EMR. As alterações são referentes à versão 5.28.1.

Data da versão inicial: 17 de janeiro de 2020

Atualizações

- Atualizado para AWS SDK para Java a versão 1.11.682
- Atualizado o Hive para a versão 2.3.6
- Atualizado o Flink para a versão 1.9.1
- Atualizado o EMRFS para a versão 2.38.0
- Atualizado o conector do DynamoDB para EMR, versão 4.13.0

Alterações, melhorias e problemas resolvidos

- Spark
  - Otimizações do desempenho do Spark
- EMRFS
  - O Guia de gerenciamento é atualizado para as configurações padrão emrfs-site.xml para uma visualização consistente.

Problemas conhecidos

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

Solução:

- SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
- Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versões de componente 5.29.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`*

começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.2.6	SDK do Amazon SageMaker Spark
emr-ddb	4.13.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.12.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.13.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.4.0	EMR S3Select Connector
emrfs	2.38.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.9.1	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.5-amzn-5	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.5-amzn-5	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.5-amzn-5	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.5-amzn-5	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	2.8.5-amzn-5	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	2.8.5-amzn-5	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.5-amzn-5	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider

Componente	Versão	Descrição
hadoop-mapred	2.8.5-amzn-5	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.5-amzn-5	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-5	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hadoop-yarn-timeline-server	2.8.5-amzn-5	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.4.10	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.10	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.10	HBase cliente de linha de comando.
hbase-rest-server	1.4.10	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.10	Serviço que fornece um endpoint Thrift para. HBase

Componente	Versão	Descrição
hcatalog-client	2.3.6-amzn-1	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.6-amzn-1	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.3.6-amzn-1	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.6-amzn-1	O cliente da linha de comando do Hive.
hive-hbase	2.3.6-amzn-1	Cliente do Hive-hbase.
hive-metastore-server	2.3.6-amzn-1	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.6-amzn-1	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.5.0-incubating	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.5.0-incubating	Biblioteca de pacotes para executar o Presto com o Hudi.

Componente	Versão	Descrição
hue-server	4.4.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	1.0.0	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.6.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.5.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.1.0	O cliente da linha de comando do Oozie.
oozie-server	5.1.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computaci onal de Código Aberto.

Componente	Versão	Descrição
phoenix-library	4.14.3- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.3- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.227	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.227	O serviço que executa partes de uma consulta.
presto-client	0.227	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística
spark-client	2.4.4	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	2.4.4	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.4	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.4.4	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	1.14.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.8.2	O notebook baseado na web que permite a análise de dados interativa.
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.

Componente	Versão	Descrição
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.29.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.29.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.

Classificações	Descrição
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.

Classificações	Descrição
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.

Classificações	Descrição
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.

Classificações	Descrição
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.

Classificações	Descrição
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.

Classificações	Descrição
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.
recordserver-env	Altere os valores no ambiente do EMR. RecordServer
recordserver-conf	Altere os valores no arquivo erver.properties RecordServer do EMR.
recordserver-log4j	Altere os valores no arquivo log4j.properties RecordServer do EMR.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.

Classificações	Descrição
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.28.1 do Amazon EMR

Versões 5.28.1 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterHub, LivyMXNet, Mahout, Oozie, Phoenix, PigPresto, Spark, Sqoop, TensorFlow, TezZeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.28.1	emr-5.28.0	emr-5.27.1	emr-5.27.0
AWS SDK for Java	1.11.659	1.11.659	1.11.615	1.11.615
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.9.0	1.9.0	1.8.1	1.8.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.10	1.4.10	1.4.10

	emr-5.28.1	emr-5.28.0	emr-5.27.1	emr-5.27.0
HCatalog	2.3.6	2.3.6	2.3.5	2.3.5
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.6	2.3.5	2.3.5
Hudi	0.5.0-incubating	0.5.0-incubating	-	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.0.0	1.0.0	1.0.0	1.0.0
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.5.1	1.5.1	1.4.0	1.4.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.3	4.14.3	4.14.2	4.14.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.227	0.227	0.224	0.224
Spark	2.4.4	2.4.4	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-5.28.1	emr-5.28.0	emr-5.27.1	emr-5.27.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas da versão 5.28.1

As notas da versão a seguir incluem informações para a versão 5.28.1 do Amazon EMR. As alterações são referentes à versão 5.28.0.

Data da versão inicial: 10 de janeiro de 2020

Alterações, melhorias e problemas resolvidos

- Spark
  - Correção de problemas de compatibilidade do Spark.
- CloudWatch Métricas
  - Foi corrigida a publicação do Amazon CloudWatch Metrics em um cluster do EMR com vários nós primários.
- Desabilitada mensagem de log
  - Desabilitada mensagem de log falsa, “... uso de versão antiga (<4.5.8) do cliente Apache http”.

Problemas conhecidos

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

Solução:

- SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
- Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versões de componente 5.28.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.2.6	SDK do Amazon SageMaker Spark
emr-ddb	4.12.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.11.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.13.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.3.0	EMR S3Select Connector
emrfs	2.37.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.9.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.5-amzn-5	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.5-amzn-5	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.5-amzn-5	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.5-amzn-5	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	2.8.5-amzn-5	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	2.8.5-amzn-5	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.5-amzn-5	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.8.5-amzn-5	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.

Componente	Versão	Descrição
hadoop-yarn-nodemanager	2.8.5-amzn-5	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-5	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hadoop-yarn-timeline-server	2.8.5-amzn-5	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.4.10	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.10	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.10	HBase cliente de linha de comando.
hbase-rest-server	1.4.10	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.10	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.6-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.

Componente	Versão	Descrição
hcatalog-server	2.3.6-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.3.6-amzn-0	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.6-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.3.6-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.3.6-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.6-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.5.0-incubating	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.5.0-incubating	Biblioteca de pacotes para executar o Presto com o Hudi.
hue-server	4.4.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop

Componente	Versão	Descrição
jupyterhub	1.0.0	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.6.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.5.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.1.0	O cliente da linha de comando do Oozie.
oozie-server	5.1.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.3- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente

Componente	Versão	Descrição
phoenix-query-server	4.14.3- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.227	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.227	O serviço que executa partes de uma consulta.
presto-client	0.227	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística
spark-client	2.4.4	Os clientes da linha de comando do Spark.
spark-history-server	2.4.4	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.

Componente	Versão	Descrição
spark-on-yarn	2.4.4	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.4.4	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	1.14.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.8.2	O notebook baseado na web que permite a análise de dados interativa.
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.28.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.28.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop

Classificações	Descrição
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.

Classificações	Descrição
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.

Classificações	Descrição
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.

Classificações	Descrição
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.

Classificações	Descrição
presto-connector-cassandra	Alterar os valores no arquivo cassandra .properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-memory	Alterar os valores no arquivo memory.pr operties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.p roperties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.propert ies do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.

Classificações	Descrição
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.
recordserver-env	Altere os valores no ambiente do EMR. RecordServer
recordserver-conf	Altere os valores no arquivo erver.properties RecordServer do EMR.
recordserver-log4j	Altere os valores no arquivo log4j.properties RecordServer do EMR.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.

Classificações	Descrição
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.28.0 do Amazon EMR

Versões 5.28.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterHub, LivyMXNet, Mahout, Oozie, Phoenix, PigPresto, Spark, Sqoop, TensorFlow, TezZeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.28.0	emr-5.27.1	emr-5.27.0	emr-5.26.0
AWS SDK for Java	1.11.659	1.11.615	1.11.615	1.11.595
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.9.0	1.8.1	1.8.1	1.8.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.10	1.4.10	1.4.10
HCatalog	2.3.6	2.3.5	2.3.5	2.3.5
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.5	2.3.5	2.3.5
Hudi	0.5.0-incubating	-	-	-

	emr-5.28.0	emr-5.27.1	emr-5.27.0	emr-5.26.0
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.0.0	1.0.0	1.0.0	0.9.6
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.5.1	1.4.0	1.4.0	1.4.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.3	4.14.2	4.14.2	4.14.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.227	0.224	0.224	0.220
Spark	2.4.4	2.4.4	2.4.4	2.4.3
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.13.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas da versão 5.28.0

As notas da versão a seguir incluem informações para a versão 5.28.0 do Amazon EMR. As alterações são referentes à 5.27.0.

Data da versão inicial: 12 de novembro de 2019

Atualizações

- Flink atualizado para a versão 1.9.0
- Atualizado o Hive para a versão 2.3.6
- Atualizado MXNet para a versão 1.5.1
- Phoenix atualizado para a versão 4.14.3
- Presto atualizado para a versão 0.227
- Zeppelin atualizado para a versão 0.8.2

Novos recursos

- O Apache Hudi agora está disponível para o Amazon EMR instalar na criação de um cluster. Para obter mais informações, consulte Hudi.
- (25 de novembro de 2019) Agora você pode optar por executar várias etapas em paralelo para melhorar a utilização do cluster e economizar custos. Pode também cancelar etapas pendentes e em execução. Para obter mais informações, consulte Trabalhar com etapas usando o console AWS CLI e.
- (3 de dezembro de 2019) Agora você pode criar e executar clusters do EMR no. AWS Outposts AWS Outposts habilita AWS serviços, infraestrutura e modelos operacionais nativos em instalações locais. Em AWS Outposts ambientes, você pode usar as mesmas AWS APIs ferramentas e infraestrutura que usa na AWS nuvem. Para obter mais informações, consulte Clusters do EMR ativados. AWS Outposts
- (11 de março de 2020) A partir da versão 5.28.0 do Amazon EMR, você pode criar e executar clusters do Amazon EMR em uma sub-rede de Zonas AWS Locais como uma extensão lógica de uma região que suporta Zonas Locais. AWS Uma zona local permite que os recursos do Amazon EMR e um subconjunto de AWS serviços, como serviços de computação e armazenamento, estejam localizados mais perto dos usuários, fornecendo acesso de latência muito baixa aos aplicativos executados localmente. Para obter uma lista das zonas locais disponíveis, consulte

Zonas locais da AWS. Para obter informações sobre como acessar as Zonas AWS Locais disponíveis, consulte Regiões, Zonas de Disponibilidade e Zonas Locais.

Atualmente, as zonas locais não são compatíveis com os Cadernos do Amazon EMR nem com conexões diretas com o Amazon EMR usando o endpoint da VPC da interface (AWS PrivateLink).

Alterações, melhorias e problemas resolvidos

- Suporte expandido do aplicativo para clusters de alta disponibilidade
  - Para obter mais informações, consulte Supported applications in an EMR cluster with Multiple Primary Nodes no Guia de gerenciamento do Amazon EMR.
- Spark
  - Otimizações da performance
- Hive
  - Otimizações da performance
- Presto
  - Otimizações da performance

Problemas conhecidos

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

  Solução:

  - SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
  - Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versões de componente 5.28.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.2.6	SDK do Amazon SageMaker Spark
emr-ddb	4.12.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.

Componente	Versão	Descrição
emr-goodies	2.11.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.13.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.3.0	EMR S3Select Connector
emrfs	2.37.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.9.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.

Componente	Versão	Descrição
hadoop-client	2.8.5-amzn-5	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.5-amzn-5	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.5-amzn-5	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.5-amzn-5	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	2.8.5-amzn-5	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	2.8.5-amzn-5	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.5-amzn-5	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.8.5-amzn-5	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.5-amzn-5	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.8.5-amzn-5	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.5-amzn-5	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.4.10	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.10	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.10	HBase cliente de linha de comando.
hbase-rest-server	1.4.10	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.10	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.6-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.6-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	2.3.6-amzn-0	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.6-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.3.6-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.3.6-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.6-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hudi	0.5.0-incubating	Estrutura de processamento incremental para alimentar o pipeline de dados com baixa latência e alta eficiência.
hudi-presto	0.5.0-incubating	Biblioteca de pacotes para executar o Presto com o Hudi.
hue-server	4.4.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	1.0.0	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.6.0-incubating	Interface REST para interagir com o Apache Spark

Componente	Versão	Descrição
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.5.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.1.0	O cliente da linha de comando do Oozie.
oozie-server	5.1.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.3- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.3- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica

Componente	Versão	Descrição
presto-coordinator	0.227	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.227	O serviço que executa partes de uma consulta.
presto-client	0.227	Cliente de linha de comando do Presto instalado em mestres de espera de um cluster de alta disponibilidade em que o servidor do Presto não é iniciado.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística
spark-client	2.4.4	Os clientes da linha de comando do Spark.
spark-history-server	2.4.4	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.4	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.4.4	As bibliotecas do Apache Spark necessárias para subordinados do YARN.

Componente	Versão	Descrição
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	1.14.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.8.2	O notebook baseado na web que permite a análise de dados interativa.
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.28.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.28.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase

Classificações	Descrição
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive- site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.pro perties do Hive.

Classificações	Descrição
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.

Classificações	Descrição
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.

Classificações	Descrição
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra .properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.

Classificações	Descrição
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.

Classificações	Descrição
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.
recordserver-env	Altere os valores no ambiente do EMR. RecordServer
recordserver-conf	Altere os valores no arquivo erver.properties RecordServer do EMR.
recordserver-log4j	Altere os valores no arquivo log4j.properties RecordServer do EMR.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.

Classificações	Descrição
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.27.1 do Amazon EMR

Versões 5.27.1 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hue, JupyterHub, Livy, MXNetMahout, Oozie, Phoenix, Pig, PrestoSpark, Sqoop, TensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR

- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.27.1	emr-5.27.0	emr-5.26.0	emr-5.25.0
AWS SDK for Java	1.11.615	1.11.615	1.11.595	1.11.566
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.1	1.8.1	1.8.0	1.8.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.10	1.4.10	1.4.9
HCatalog	2.3.5	2.3.5	2.3.5	2.3.5
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.5	2.3.5	2.3.5	2.3.5
Hudi	-	-	-	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-5.27.1	emr-5.27.0	emr-5.26.0	emr-5.25.0
JupyterHub	1.0.0	1.0.0	0.9.6	0.9.6
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.4.0	1.4.0	1.4.0	1.4.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.2	4.14.2	4.14.2	4.14.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.224	0.224	0.220	0.220
Spark	2.4.4	2.4.4	2.4.3	2.4.3
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.13.1	1.13.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas da versão 5.27.1

Esta é uma versão de patch. Todas as aplicações e componentes são os mesmos que os da versão anterior do Amazon EMR.

Status de suporte do Instance Metadata Service (IMDS) V2: os componentes do Amazon EMR 5.23.1, 5.27.1 e 5.32 ou posteriores são usados para todas as chamadas do IMDS. IMDSv2 Para chamadas IMDS no código do aplicativo, você pode usar ambos IMDSv1 e IMDSv2, ou configurar

o IMDS para uso somente IMDSv2 para aumentar a segurança. Para outras versões 5.x do EMR, a IMDSv1 desativação causa falha na inicialização do cluster.

Versões de componente 5.27.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.2.4	SDK do Amazon SageMaker Spark
emr-ddb	4.12.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.11.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.13.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.3.0	EMR S3Select Connector

Componente	Versão	Descrição
emrfs	2.36.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.8.1	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.5-amzn-4	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.5-amzn-4	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.5-amzn-4	O cliente de linha de comando e biblioteca do HDFS

Componente	Versão	Descrição
hadoop-hdfs-namenode	2.8.5-amzn-4	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	2.8.5-amzn-4	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	2.8.5-amzn-4	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.5-amzn-4	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-4	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.5-amzn-4	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.8.5-amzn-4	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.5-amzn-4	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.

Componente	Versão	Descrição
hbase-hmaster	1.4.10	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.10	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.10	HBase cliente de linha de comando.
hbase-rest-server	1.4.10	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.10	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.5-amzn-1	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.5-amzn-1	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.3.5-amzn-1	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.5-amzn-1	O cliente da linha de comando do Hive.
hive-hbase	2.3.5-amzn-1	Cliente do Hive-hbase.

Componente	Versão	Descrição
hive-metastore-server	2.3.5-amzn-1	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.5-amzn-1	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.4.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	1.0.0	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.6.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.4.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA

Componente	Versão	Descrição
oozie-client	5.1.0	O cliente da linha de comando do Oozie.
oozie-server	5.1.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.2- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.2- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.224	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.224	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística
spark-client	2.4.4	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	2.4.4	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.4	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.4.4	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	1.14.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.8.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.

Componente	Versão	Descrição
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.27.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.27.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.

Classificações	Descrição
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.

Classificações	Descrição
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.

Classificações	Descrição
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.

Classificações	Descrição
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.

Classificações	Descrição
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.

Classificações	Descrição
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.
recordserver-env	Altere os valores no ambiente do EMR. RecordServer
recordserver-conf	Altere os valores no arquivo erver.properties RecordServer do EMR.
recordserver-log4j	Altere os valores no arquivo log4j.properties RecordServer do EMR.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.

Classificações	Descrição
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.27.0 do Amazon EMR

Versões 5.27.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hue, JupyterHub, Livy, MXNetMahout, Oozie, Phoenix, Pig, PrestoSpark, Sqoop, TensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.27.0	emr-5.26.0	emr-5.25.0	emr-5.24.1
AWS SDK for Java	1.11.615	1.11.595	1.11.566	1.11.546
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.1	1.8.0	1.8.0	1.8.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.10	1.4.9	1.4.9

	emr-5.27.0	emr-5.26.0	emr-5.25.0	emr-5.24.1
HCatalog	2.3.5	2.3.5	2.3.5	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.5	2.3.5	2.3.5	2.3.4
Hudi	-	-	-	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.0.0	0.9.6	0.9.6	0.9.6
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.4.0	1.4.0	1.4.0	1.4.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.2	4.14.2	4.14.1	4.14.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.224	0.220	0.220	0.219
Spark	2.4.4	2.4.3	2.4.3	2.4.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.13.1	1.13.1	1.12.0
Tez	0.9.2	0.9.2	0.9.2	0.9.1

	emr-5.27.0	emr-5.26.0	emr-5.25.0	emr-5.24.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.13

Notas da versão 5.27.0

As notas da versão a seguir incluem informações para a versão 5.27.0 do Amazon EMR. As alterações são referentes à versão 5.26.0.

Data da versão inicial: 23 de setembro de 2019

Atualizações

- AWS SDK para Java 1.11.615
- Flink 1.8.1
- JupyterHub 1.0.0
- Spark 2.4.4
- Tensorflow 1.14.0
- Conectores e drivers:
  - Conector do DynamoDB 4.12.0

Novos recursos

- (24 de outubro de 2019) Os seguintes novos atributos dos Cadernos do EMR estão disponíveis em todas as versões do Amazon EMR.
  - É possível associar repositórios do Git a Cadernos do EMR para armazenar os cadernos em um ambiente controlado de versão. Você pode compartilhar códigos com pares e reutilizar cadernos Jupyter existentes por meio de repositórios do Git remotos. Para obter mais informações, consulte Associate Git Repositories with Amazon EMR Notebooks no Guia de gerenciamento do Amazon EMR.

  - O utilitário nbdime agora está disponível em Cadernos do EMR para simplificar a comparação e a mesclagem de cadernos.
  - Agora há suporte para notebooks EMR. JupyterLab JupyterLab é um ambiente de desenvolvimento interativo baseado na Web totalmente compatível com os notebooks Jupyter. Agora você pode optar por abrir seu caderno em qualquer um dos editores de cadernos JupyterLab ou no editor de cadernos Jupyter.
- (30 de outubro de 2019) Com as versões 5.25.0 e posteriores do Amazon EMR, é possível conectar-se ao servidor de histórico do Spark na página Resumo do cluster ou na guia Histórico da aplicação no console. Em vez de configurar um proxy da Web por meio de uma conexão SSH, você pode acessar rapidamente a interface do usuário do servidor de histórico do Spark para visualizar as métricas da aplicação e acessar arquivos de log relevantes para clusters ativos e encerrados. Para obter mais informações, consulte Acesso fora do cluster a interfaces de usuário de aplicações persistentes no Guia de gerenciamento do Amazon EMR.

Alterações, melhorias e problemas resolvidos

- Cluster do Amazon EMR com vários nós primários
  - Você pode instalar e executar o Flink em um cluster do Amazon EMR com vários nós primários. Para obter mais informações, consulte Supported applications and features.
  - Você pode configurar a criptografia transparente do HDFS em um cluster do Amazon EMR com vários nós primários. Para obter mais informações, consulte Criptografia transparente do HDFS em clusters do EMR com vários nós primários.
  - Agora você pode modificar a configuração das aplicações em execução em um cluster do Amazon EMR com vários nós primários. Para obter mais informações, consulte Supplying a Configuration for an Instance Group in a Running Cluster.
- Conector Amazon EMR-DynamoDB
  - O conector Amazon EMR-DynamoDB agora é compatível com os seguintes tipos de dados do DynamoDB: booleanos, lista, mapa, item, nulos. Para obter mais informações, consulte Configurar uma tabela do Hive para executar comandos do Hive.

Problemas conhecidos

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

Solução:

- SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
- Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versões de componente 5.27.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`*

começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.2.4	SDK do Amazon SageMaker Spark
emr-ddb	4.12.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.11.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.13.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.3.0	EMR S3Select Connector
emrfs	2.36.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.8.1	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.5-amzn-4	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.5-amzn-4	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.5-amzn-4	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.5-amzn-4	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	2.8.5-amzn-4	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	2.8.5-amzn-4	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.5-amzn-4	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider

Componente	Versão	Descrição
hadoop-mapred	2.8.5-amzn-4	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.5-amzn-4	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-4	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hadoop-yarn-timeline-server	2.8.5-amzn-4	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.4.10	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.10	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.10	HBase cliente de linha de comando.
hbase-rest-server	1.4.10	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.10	Serviço que fornece um endpoint Thrift para. HBase

Componente	Versão	Descrição
hcatalog-client	2.3.5-amzn-1	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.5-amzn-1	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.3.5-amzn-1	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.5-amzn-1	O cliente da linha de comando do Hive.
hive-hbase	2.3.5-amzn-1	Cliente do Hive-hbase.
hive-metastore-server	2.3.5-amzn-1	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.5-amzn-1	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.4.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	1.0.0	Servidor de vários usuários para blocos de anotações Jupyter

Componente	Versão	Descrição
livy-server	0.6.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.4.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.1.0	O cliente da linha de comando do Oozie.
oozie-server	5.1.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.2- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.2- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica

Componente	Versão	Descrição
presto-coordinator	0.224	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.224	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística
spark-client	2.4.4	Os clientes da linha de comando do Spark.
spark-history-server	2.4.4	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.4	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.4.4	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	1.14.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.

Componente	Versão	Descrição
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.8.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.27.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.27.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.

Classificações	Descrição
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.

Classificações	Descrição
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.

Classificações	Descrição
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.

Classificações	Descrição
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.

Classificações	Descrição
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.

Classificações	Descrição
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.

Classificações	Descrição
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.
recordserver-env	Altere os valores no ambiente do EMR. RecordServer
recordserver-conf	Altere os valores no arquivo erver.properties RecordServer do EMR.
recordserver-log4j	Altere os valores no arquivo log4j.properties RecordServer do EMR.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.con f do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.

Classificações	Descrição
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.26.0 do Amazon EMR

Versões 5.26.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hue, JupyterHub, Livy, MXNetMahout, Oozie, Phoenix, Pig, PrestoSpark, Sqoop, TensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.26.0	emr-5.25.0	emr-5.24.1	emr-5.24.0
AWS SDK for Java	1.11.595	1.11.566	1.11.546	1.11.546
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.0	1.8.0	1.8.0	1.8.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.9	1.4.9	1.4.9
HCatalog	2.3.5	2.3.5	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.5	2.3.5	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.6	0.9.6	0.9.6	0.9.6
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.4.0	1.4.0	1.4.0	1.4.0

	emr-5.26.0	emr-5.25.0	emr-5.24.1	emr-5.24.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.2	4.14.1	4.14.1	4.14.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.220	0.220	0.219	0.219
Spark	2.4.3	2.4.3	2.4.2	2.4.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.13.1	1.13.1	1.12.0	1.12.0
Tez	0.9.2	0.9.2	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.13	3.4.13

Notas da versão 5.26.0

As notas da versão a seguir incluem informações para a versão 5.26.0 do Amazon EMR. As alterações são referentes à versão 5.25.0.

Data da versão inicial: 8 de agosto de 2019

Data da última atualização: 19 de agosto de 2019

Atualizações

- AWS SDK para Java 1.11.595
- HBase 1.4.10
- Phoenix 4.14.2

- Conectores e drivers:
  - Conector do DynamoDB 4.11.0
  - Conector do MariaDB 2.4.2
  - Driver JDBC do Amazon Redshift, 1.2.32.1056

Novos recursos

- (Beta) Com a versão 5.26.0 do Amazon EMR, você pode iniciar um cluster que se integre ao Lake Formation. Essa integração fornece acesso refinado em nível de coluna a bancos de dados e tabelas no Glue Data Catalog. AWS Ela também permite logon único federado para Cadernos do EMR ou para o Apache Zeppelin em um sistema de identidade empresarial. Para obter mais informações, consulte Integração do Amazon EMR AWS Lake Formation com (Beta).
- (19 de agosto de 2019) O bloqueio de acesso público do Amazon EMR agora está disponível em todas as versões do Amazon EMR compatíveis com grupos de segurança. Bloquear o acesso público é uma configuração de toda a conta aplicada a cada AWS região. Bloquear o acesso público impede que um cluster seja iniciado quando qualquer grupo de segurança associado ao cluster tem uma regra que permite tráfego de entrada de IPv4 0.0.0.0/0 ou IPv6 : :/0 (acesso público) em uma porta, a menos que uma porta seja especificada como uma exceção. A porta 22 é uma exceção por padrão. Para obter mais informações, consulte Usar o bloqueio de acesso público do Amazon EMR no Guia de gerenciamento do Amazon EMR.

Alterações, melhorias e problemas resolvidos

- Cadernos do EMR
  - Com as versões 5.26.0 e posteriores do EMR, os Cadernos do EMR são compatíveis com as bibliotecas Python com escopo de caderno, além das bibliotecas Python padrão. Você pode instalar bibliotecas com escopo de caderno de dentro do editor de caderno sem precisar recriar um cluster ou reanexar um caderno a um cluster. As bibliotecas com escopo de caderno são criadas em um ambiente Python virtual para serem aplicadas somente à sessão de caderno atual. Isso permite isolar as dependências do caderno. Para obter mais informações, consulte Usar bibliotecas com escopo de caderno no Guia de gerenciamento do Amazon EMR.
- EMRFS
  - Você pode ativar um recurso ETag de verificação (Beta) configurando `fs.s3.consistent.metadata.etag.verification.enabled` como `true`. Com esse recurso, o EMRFS usa o Amazon ETags S3 para verificar se os objetos que estão sendo lidos

são a versão mais recente disponível. Esse recurso é útil para casos de read-after-update uso em que os arquivos no Amazon S3 são sobrescritos, mantendo o mesmo nome. Atualmente, esse recurso de ETag verificação não funciona com o S3 Select. Para obter mais informações, consulte Configurar visualização consistente.

- Spark
  - As seguintes otimizações agora estão habilitadas por padrão: remoção dinâmica de partições, DISTINCT antes de INTERSECT, melhorias na inferência de estatísticas do plano SQL para JOIN seguida por consultas DISTINCT, nivelamento de subconsultas escalares, reordenamento otimizado de junções e junção com filtro de Bloom. Para obter mais informações, consulte Otimizar a performance do Spark.
  - Aprimorada a geração de código de estágio completo para Sort-Merge Join.
  - Aprimorado o fragmento de consulta e a reutilização de subconsultas.
  - Melhorias na pré-alocação de executores na inicialização do Spark.
  - As junções com filtro de Bloom não são mais aplicadas quando o lado menor da junção inclui uma dica de transmissão.
- Tez
  - Resolvido um problema com o Tez. A IU do Tez agora funciona em um cluster do Amazon EMR com vários nós primários.

Problemas conhecidos

- Os recursos aprimorados de geração de código em todo o estágio de Sort Merge Join podem aumentar a pressão de memória quando habilitados. Essa otimização melhora a performance, mas pode resultar em novas tentativas ou falhas de trabalho se `spark.yarn.executor.memoryOverheadFactor` não for ajustado para fornecer memória suficiente. Para desabilitar esse atributo, defina `spark.sql.sortMergeJoinExec.extendedCodegen.enabled` como falso.
- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução

automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

Solução:

- SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
- Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versões de componente 5.26.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.2.4	SDK do Amazon SageMaker Spark
emr-ddb	4.11.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.10.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.12.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.3.0	EMR S3Select Connector
emrfs	2.35.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.8.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.5-amzn-4	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.5-amzn-4	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.5-amzn-4	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.5-amzn-4	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	2.8.5-amzn-4	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	2.8.5-amzn-4	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.5-amzn-4	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.8.5-amzn-4	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.

Componente	Versão	Descrição
hadoop-yarn-nodemanager	2.8.5-amzn-4	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-4	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hadoop-yarn-timeline-server	2.8.5-amzn-4	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.4.10	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.10	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.10	HBase cliente de linha de comando.
hbase-rest-server	1.4.10	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.10	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.5-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.

Componente	Versão	Descrição
hcatalog-server	2.3.5-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.3.5-amzn-0	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.5-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.3.5-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.3.5-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.5-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.4.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	0.9.6	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.6.0-incubating	Interface REST para interagir com o Apache Spark

Componente	Versão	Descrição
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.4.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.1.0	O cliente da linha de comando do Oozie.
oozie-server	5.1.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.2- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.2- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica

Componente	Versão	Descrição
presto-coordinator	0.220	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.220	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística
spark-client	2.4.3	Os clientes da linha de comando do Spark.
spark-history-server	2.4.3	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.3	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.4.3	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	1.13.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.

Componente	Versão	Descrição
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.8.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.26.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.26.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.

Classificações	Descrição
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.

Classificações	Descrição
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.

Classificações	Descrição
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.

Classificações	Descrição
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.

Classificações	Descrição
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.

Classificações	Descrição
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
recordserver-env	Altere os valores no ambiente do EMR. RecordServer
recordserver-conf	Altere os valores no arquivo erver.properties RecordServer do EMR.
recordserver-log4j	Altere os valores no arquivo log4j.properties RecordServer do EMR.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.

Classificações	Descrição
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.25.0 do Amazon EMR

Versões 5.25.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hue, JupyterHub, Livy, MXNetMahout, Oozie, Phoenix, Pig, PrestoSpark, Sqoop, TensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.25.0	emr-5.24.1	emr-5.24.0	emr-5.23.1
AWS SDK for Java	1.11.566	1.11.546	1.11.546	1.11.519
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.0	1.8.0	1.8.0	1.7.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.9	1.4.9	1.4.9

	emr-5.25.0	emr-5.24.1	emr-5.24.0	emr-5.23.1
HCatalog	2.3.5	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.5	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.4.0	4.4.0	4.4.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.6	0.9.6	0.9.6	0.9.4
Livy	0.6.0	0.6.0	0.6.0	0.5.0
MXNet	1.4.0	1.4.0	1.4.0	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.1	4.14.1	4.14.1	4.14.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.220	0.219	0.219	0.215
Spark	2.4.3	2.4.2	2.4.2	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.13.1	1.12.0	1.12.0	1.12.0
Tez	0.9.2	0.9.1	0.9.1	0.9.1

	emr-5.25.0	emr-5.24.1	emr-5.24.0	emr-5.23.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.13	3.4.13	3.4.13

Notas da versão 5.25.0

As notas da versão a seguir incluem informações para a versão 5.25.0 do Amazon EMR. As alterações são referentes à versão 5.24.1.

Data da versão inicial: 17 de julho de 2019

Data da última atualização: 30 de outubro de 2019

Amazon EMR 5.25.0

Atualizações

- AWS SDK para Java 1.11.566
- Hive 2.3.5
- Presto 0.220
- Spark 2.4.3
- TensorFlow 1.13.1
- Tez 0.9.2
- Zookeeper 3.4.14

Novos recursos

- (30 de outubro de 2019) Desde a versão 5.25.0 do Amazon EMR, é possível conectar-se ao servidor de histórico do Spark na página Resumo do cluster ou na guia Histórico da aplicação no console. Em vez de configurar um proxy da Web por meio de uma conexão SSH, você pode acessar rapidamente a interface do usuário do servidor de histórico do Spark para visualizar as métricas da aplicação e acessar arquivos de log relevantes para clusters ativos e encerrados. Para

obter mais informações, consulte Acesso fora do cluster a interfaces de usuário de aplicações persistentes no Guia de gerenciamento do Amazon EMR.

Alterações, melhorias e problemas resolvidos

- Spark
  - Aprimorada a performance de algumas junções usando filtros de Bloom para pré-filtrar as entradas. A otimização é desabilitada por padrão e pode ser habilitada com a definição do parâmetro `spark.sql.bloomFilterJoin.enabled` de configuração do Spark como `true`.
  - Aprimorada a performance do agrupamento por colunas do tipo string.
  - Melhorou a memória padrão do executor Spark e a configuração dos núcleos dos tipos de instância R4 para clusters sem instalação. HBase
  - Resolvido um problema anterior com o atributo de remoção dinâmica de partições, em que a tabela removida precisava estar no lado esquerdo da junção.
  - Aprimorada a otimização do DISTINCT antes do INTERSECT para ser aplicada a casos adicionais envolvendo aliases.
  - Aprimorada a inferência de estatísticas do plano SQL para JOIN seguida por consultas DISTINCT. Essa melhoria é desabilitada por padrão e pode ser habilitada pela definição do parâmetro `spark.sql.statsImprovements.enabled` de configuração do Spark como `true`. Essa otimização é exigida pelo atributo Distinct antes do Intersect e será habilitada automaticamente quando `spark.sql.optimizer.distinctBeforeIntersect.enabled` estiver definido como `true`.
  - Otimizada a ordem de junção com base no tamanho da tabela e nos filtros. Essa otimização é desativada por padrão e pode ser ativada com a definição do parâmetro `spark.sql.optimizer.sizeBasedJoinReorder.enabled` de configuração do Spark como `true`.

  Para obter mais informações, consulte Otimizar a performance do Spark.
- EMRFS
  - A configuração do EMRFS, `fs.s3.buckets.create.enabled`, agora está desabilitada por padrão. Por meio de testes, descobrimos que a desabilitação dessa configuração melhora a performance e evita a criação não intencional de buckets do S3. Se sua aplicação depende dessa funcionalidade, você pode habilitá-la definindo a propriedade `fs.s3.buckets.create.enabled` como `true` na classificação de configuração `emrfs-site`. Para obter informações, consulte Supplying a Configuration when Creating a Cluster.

- Melhorias na criptografia de disco local e na criptografia do S3 nas configurações de segurança (5 de agosto de 2019)
  - Separadas as configurações de criptografia do Amazon S3 das configurações de criptografia de disco local na configuração de segurança.
  - Adicionada uma opção para habilitar a criptografia do EBS com as versões 5.24.0 e posteriores. Selecionar essa opção criptografa o volume do dispositivo raiz, além dos volumes de armazenamento. As versões anteriores exigiam o uso de uma AMI personalizada para criptografar o volume do dispositivo raiz.
  - Para obter mais informações, consulte Opções de criptografia no Guia de gerenciamento do Amazon EMR.

Problemas conhecidos

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

  Solução:

  - SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
  - Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas

definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versões de componente 5.25.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.2.4	SDK do Amazon SageMaker Spark
emr-ddb	4.10.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.9.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.11.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.

Componente	Versão	Descrição
emr-s3-select	1.3.0	EMR S3Select Connector
emrfs	2.34.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.8.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.5-amzn-4	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.5-amzn-4	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.5-amzn-4	O cliente de linha de comando e biblioteca do HDFS

Componente	Versão	Descrição
hadoop-hdfs-namenode	2.8.5-amzn-4	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	2.8.5-amzn-4	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	2.8.5-amzn-4	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.5-amzn-4	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-4	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.5-amzn-4	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.8.5-amzn-4	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.5-amzn-4	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.

Componente	Versão	Descrição
hbase-hmaster	1.4.9	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.9	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.9	HBase cliente de linha de comando.
hbase-rest-server	1.4.9	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.9	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.5-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.5-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.3.5-amzn-0	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.5-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.3.5-amzn-0	Cliente do Hive-hbase.

Componente	Versão	Descrição
hive-metastore-server	2.3.5-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.5-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.4.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	0.9.6	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.6.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.4.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA

Componente	Versão	Descrição
oozie-client	5.1.0	O cliente da linha de comando do Oozie.
oozie-server	5.1.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.1- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.1- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.220	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.220	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística
spark-client	2.4.3	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	2.4.3	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.3	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.4.3	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	1.13.1	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.2	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.8.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.14	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.

Componente	Versão	Descrição
zookeeper-client	3.4.14	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.25.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.25.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.

Classificações	Descrição
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.

Classificações	Descrição
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.

Classificações	Descrição
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.

Classificações	Descrição
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.

Classificações	Descrição
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.

Classificações	Descrição
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
recordserver-env	Altere os valores no ambiente do EMR. RecordServer
recordserver-conf	Altere os valores no arquivo erver.properties RecordServer do EMR.
recordserver-log4j	Altere os valores no arquivo log4j.properties RecordServer do EMR.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.con f do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.

Classificações	Descrição
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.24.1 do Amazon EMR

Versões 5.24.1 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hue, JupyterHub, Livy, MXNetMahout, Oozie, Phoenix, Pig, PrestoSpark, Sqoop, TensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR

- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.24.1	emr-5.24.0	emr-5.23.1	emr-5.23.0
AWS SDK for Java	1.11.546	1.11.546	1.11.519	1.11.519
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.0	1.8.0	1.7.1	1.7.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.9	1.4.9	1.4.9
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.4.0	4.4.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-5.24.1	emr-5.24.0	emr-5.23.1	emr-5.23.0
JupyterHub	0.9.6	0.9.6	0.9.4	0.9.4
Livy	0.6.0	0.6.0	0.5.0	0.5.0
MXNet	1.4.0	1.4.0	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.1	4.14.1	4.14.1	4.14.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.219	0.219	0.215	0.215
Spark	2.4.2	2.4.2	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

Notas da versão 5.24.1

As notas da versão a seguir incluem informações para a versão 5.24.1 do Amazon EMR. As alterações são referentes à versão 5.24.0.

Data da versão inicial: 26 de junho de 2019

Alterações, melhorias e problemas resolvidos

- Atualizada a AMI padrão do Amazon Linux para Amazon EMR para incluir atualizações de segurança importantes do kernel Linux, incluindo o problema de negação de serviço do TCP SACK (AWS-2019-005).

Problemas conhecidos

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

  Solução:

  - SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
  - Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versões de componente 5.24.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.2.1	SDK do Amazon SageMaker Spark
emr-ddb	4.9.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.8.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.11.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.3.0	EMR S3Select Connector
emrfs	2.33.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.

Componente	Versão	Descrição
flink-client	1.8.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.5-amzn-4	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.5-amzn-4	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.5-amzn-4	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.5-amzn-4	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.

Componente	Versão	Descrição
hadoop-hdfs-journalnode	2.8.5-amzn-4	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	2.8.5-amzn-4	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.5-amzn-4	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-4	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.5-amzn-4	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.8.5-amzn-4	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.5-amzn-4	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.4.9	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.

Componente	Versão	Descrição
hbase-region-server	1.4.9	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.9	HBase cliente de linha de comando.
hbase-rest-server	1.4.9	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.9	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.4-amzn-2	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.4-amzn-2	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.3.4-amzn-2	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.4-amzn-2	O cliente da linha de comando do Hive.
hive-hbase	2.3.4-amzn-2	Cliente do Hive-hbase.
hive-metastore-server	2.3.4-amzn-2	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.

Componente	Versão	Descrição
hive-server2	2.3.4-amzn-2	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.4.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	0.9.6	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.6.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.4.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.1.0	O cliente da linha de comando do Oozie.
oozie-server	5.1.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.

Componente	Versão	Descrição
opencv	3.4.0	Biblioteca de Visão Computacional de Código Aberto.
phoenix-library	4.14.1- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.1- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.219	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.219	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística
spark-client	2.4.2	Os clientes da linha de comando do Spark.
spark-history-server	2.4.2	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.2	O mecanismo de execução em memória para o YARN.

Componente	Versão	Descrição
spark-yarn-slave	2.4.2	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	1.12.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.1	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.8.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.13	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.13	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.24.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.24.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase

Classificações	Descrição
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive- site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.pro perties do Hive.

Classificações	Descrição
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.

Classificações	Descrição
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.

Classificações	Descrição
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra .properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.

Classificações	Descrição
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.

Classificações	Descrição
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.24.0 do Amazon EMR

Versões 5.24.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hue, JupyterHub, Livy, MXNetMahout, Oozie, Phoenix, Pig, PrestoSpark, Sqoop, TensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.24.0	emr-5.23.1	emr-5.23.0	emr-5.22.0
AWS SDK for Java	1.11.546	1.11.519	1.11.519	1.11.510
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.0	1.7.1	1.7.1	1.7.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.9	1.4.9	1.4.9

	emr-5.24.0	emr-5.23.1	emr-5.23.0	emr-5.22.0
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.4.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.6	0.9.4	0.9.4	0.9.4
Livy	0.6.0	0.5.0	0.5.0	0.5.0
MXNet	1.4.0	1.3.1	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.1	4.14.1	4.14.1	4.14.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.219	0.215	0.215	0.215
Spark	2.4.2	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1

	emr-5.24.0	emr-5.23.1	emr-5.23.0	emr-5.22.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

Notas da versão 5.24.0

As notas da versão a seguir incluem informações para a versão 5.24.0 do Amazon EMR. As alterações são referentes à versão 5.23.0.

Data da versão inicial: 11 de junho de 2019

Data da última atualização: 5 de agosto de 2019

Atualizações

- Flink 1.8.0
- Hue 4.4.0
- JupyterHub 0.9.6
- Livy 0.6.0
- MxNet 1.4.0
- Presto 0.219
- Spark 2.4.2
- AWS SDK para Java 1.11.546
- Conectores e drivers:
  - Conector do DynamoDB 4.9.0
  - Conector do MariaDB 2.4.1
  - Driver JDBC do Amazon Redshift, 1.2.27.1051

Alterações, melhorias e problemas resolvidos

- Spark

- Adicionada otimização para remover partições dinamicamente. A otimização está desabilitada por padrão. Para habilitá-la, defina o parâmetro `spark.sql.dynamicPartitionPruning.enabled` de configuração do Spark como `true`.
- Aprimorada a performance de consultas `INTERSECT`. Essa otimização está desabilitada por padrão. Para habilitá-la, defina o parâmetro `spark.sql.optimizer.distinctBeforeIntersect.enabled` de configuração do Spark como `true`.
- Adicionada otimização para nivelar subconsultas escalares com agregados que usam a mesma relação. A otimização está desabilitada por padrão. Para habilitá-la, defina o parâmetro `spark.sql.optimizer.flattenScalarSubqueriesWithAggregates.enabled` de configuração do Spark como `true`.
- Aprimorada a geração de código em todo o estágio.

  Para obter mais informações, consulte Otimizar a performance do Spark.

- Melhorias na criptografia de disco local e na criptografia do S3 nas configurações de segurança (5 de agosto de 2019)
  - Separadas as configurações de criptografia do Amazon S3 das configurações de criptografia de disco local na configuração de segurança.
  - Adicionada uma opção para habilitar a criptografia do EBS. Selecionar essa opção criptografa o volume do dispositivo raiz, além dos volumes de armazenamento. As versões anteriores exigiam o uso de uma AMI personalizada para criptografar o volume do dispositivo raiz.
  - Para obter mais informações, consulte Opções de criptografia no Guia de gerenciamento do Amazon EMR.

Problemas conhecidos

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

Solução:

- SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
- Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versões de componente 5.24.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.2.1	SDK do Amazon SageMaker Spark
emr-ddb	4.9.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.8.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.11.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.3.0	EMR S3Select Connector
emrfs	2.33.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.8.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.5-amzn-4	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.5-amzn-4	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.5-amzn-4	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.5-amzn-4	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	2.8.5-amzn-4	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	2.8.5-amzn-4	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.5-amzn-4	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.8.5-amzn-4	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.

Componente	Versão	Descrição
hadoop-yarn-nodemanager	2.8.5-amzn-4	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-4	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hadoop-yarn-timeline-server	2.8.5-amzn-4	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.4.9	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.9	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.9	HBase cliente de linha de comando.
hbase-rest-server	1.4.9	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.9	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.4-amzn-2	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.

Componente	Versão	Descrição
hcatalog-server	2.3.4-amzn-2	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.3.4-amzn-2	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.4-amzn-2	O cliente da linha de comando do Hive.
hive-hbase	2.3.4-amzn-2	Cliente do Hive-hbase.
hive-metastore-server	2.3.4-amzn-2	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.4-amzn-2	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.4.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	0.9.6	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.6.0-incubating	Interface REST para interagir com o Apache Spark

Componente	Versão	Descrição
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.4.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.1.0	O cliente da linha de comando do Oozie.
oozie-server	5.1.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.1- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.1- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica

Componente	Versão	Descrição
presto-coordinator	0.219	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.219	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística
spark-client	2.4.2	Os clientes da linha de comando do Spark.
spark-history-server	2.4.2	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.2	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.4.2	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	1.12.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.

Componente	Versão	Descrição
tez-on-yarn	0.9.1	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.8.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.13	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.13	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.24.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.24.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.

Classificações	Descrição
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.

Classificações	Descrição
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.

Classificações	Descrição
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.

Classificações	Descrição
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.

Classificações	Descrição
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.

Classificações	Descrição
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.

Classificações	Descrição
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.23.1 do Amazon EMR

Versões 5.23.1 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hue, JupyterHub, Livy, MXNetMahout, Oozie, Phoenix, Pig, PrestoSpark, Sqoop, TensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.23.1	emr-5.23.0	emr-5.22.0	emr-5.21.2
AWS SDK for Java	1.11.519	1.11.519	1.11.510	1.11.479
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.1	1.7.1	1.7.1	1.7.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.9	1.4.9	1.4.8
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-

	emr-5.23.1	emr-5.23.0	emr-5.22.0	emr-5.21.2
Hue	4.3.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.0.0
Phoenix	4.14.1	4.14.1	4.14.1	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.215	0.215	0.215
Spark	2.4.0	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

Notas da versão 5.23.1

Esta é uma versão de patch. Todas as aplicações e componentes são os mesmos que os da versão anterior do Amazon EMR.

Status de suporte do Instance Metadata Service (IMDS) V2: os componentes do Amazon EMR 5.23.1, 5.27.1 e 5.32 ou posteriores são usados para todas as chamadas do IMDS. IMDSv2 Para chamadas IMDS no código do aplicativo, você pode usar ambos IMDSv1 e IMDSv2, ou configurar o IMDS para uso somente IMDSv2 para aumentar a segurança. Para outras versões 5.x do EMR, a IMDSv1 desativação causa falha na inicialização do cluster.

Versões de componente 5.23.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.2.1	SDK do Amazon SageMaker Spark
emr-ddb	4.8.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.7.0	Bibliotecas convenientes para o ecossistema do Hadoop.

Componente	Versão	Descrição
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.11.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.2.0	EMR S3Select Connector
emrfs	2.32.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.7.1	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.5-amzn-3	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.

Componente	Versão	Descrição
hadoop-hdfs-datanode	2.8.5-amzn-3	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.5-amzn-3	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.5-amzn-3	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-hdfs-journalnode	2.8.5-amzn-3	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	2.8.5-amzn-3	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.5-amzn-3	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.8.5-amzn-3	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.5-amzn-3	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-3	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.

Componente	Versão	Descrição
hadoop-yarn-timeline-server	2.8.5-amzn-3	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.4.9	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.9	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.9	HBase cliente de linha de comando.
hbase-rest-server	1.4.9	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.9	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.4-amzn-1	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.4-amzn-1	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.3.4-amzn-1	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.

Componente	Versão	Descrição
hive-client	2.3.4-amzn-1	O cliente da linha de comando do Hive.
hive-hbase	2.3.4-amzn-1	Cliente do Hive-hbase.
hive-metastore-server	2.3.4-amzn-1	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.4-amzn-1	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.3.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	0.9.4	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.5.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.3.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.

Componente	Versão	Descrição
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.1.0	O cliente da linha de comando do Oozie.
oozie-server	5.1.0	O serviço que aceita solicitações de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computacional de Código Aberto.
phoenix-library	4.14.1- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.1- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.215	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.215	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística

Componente	Versão	Descrição
spark-client	2.4.0	Os clientes da linha de comando do Spark.
spark-history-server	2.4.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.4.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	1.12.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.1	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.8.1	O notebook baseado na web que permite um data analytics interativo.

Componente	Versão	Descrição
zookeeper-server	3.4.13	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.13	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.23.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.23.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.

Classificações	Descrição
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.

Classificações	Descrição
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.

Classificações	Descrição
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.

Classificações	Descrição
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.

Classificações	Descrição
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.

Classificações	Descrição
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.

Classificações	Descrição
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.23.0 do Amazon EMR

Versões 5.23.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hue, JupyterHub, Livy, MXNetMahout, Oozie, Phoenix, Pig, PrestoSpark, Sqoop, TensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.23.0	emr-5.22.0	emr-5.21.2	emr-5.21.1
AWS SDK for Java	1.11.519	1.11.510	1.11.479	1.11.479
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.1	1.7.1	1.7.0	1.7.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.9	1.4.8	1.4.8
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-

	emr-5.23.0	emr-5.22.0	emr-5.21.2	emr-5.21.1
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.0.0	5.0.0
Phoenix	4.14.1	4.14.1	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.215	0.215	0.215
Spark	2.4.0	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

Notas da versão 5.23.0

As notas da versão a seguir incluem informações para a versão 5.23.0 do Amazon EMR. As alterações são referentes à versão 5.22.0.

Data da versão inicial: 1.º de abril de 2019

Data da última atualização: 30 de abril de 2019

Atualizações

- AWS SDK para Java 1.11.519

Novos recursos

- (30 de abril de 2019) Com o Amazon EMR 5.23.0 e versões posteriores, você pode iniciar um cluster com três nós principais para oferecer suporte à alta disponibilidade de aplicativos como YARN Resource Manager, HDFS, Spark, Hive e NameNode Ganglia. O nó primário não é mais um possível ponto de falha único com esse recurso. Se um dos nós primários apresenta falha, o Amazon EMR executa failover automaticamente para um nó primário em espera e substitui o nó primário com falha por um novo com as mesmas ações de configuração e bootstrap. Para obter mais informações, consulte Plan and Configure Primary Nodes.

Problemas conhecidos

- Interface do usuário do Tez (corrigida na versão 5.26.0 do Amazon EMR)

  A interface do usuário do Tez não funciona em um cluster do EMR com vários nós primários.
- Hue (corrigido na versão 5.24.0 do Amazon EMR)
  - O Hue executado no Amazon EMR não é compatível com o Solr. Desde a versão 5.20.0 do Amazon EMR, um problema de configuração incorreta faz com que o Solr seja habilitado e uma mensagem de erro inofensiva semelhante à seguinte seja exibida:

    ```
    Solr server could not be contacted properly:
    HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
    port=1978): Max retries exceeded with url: /solr/admin/info/
    system?user.name=hue&doAs=administrator&wt=json (Caused by
    NewConnectionError(': Failed to establish a new connection: [Errno 111]
    Connection refused',))
    ```

    Para evitar que a mensagem de erro do Solr seja exibida:

    1. Conecte-se à linha de comando do nó primário usando SSH.
    2. Use um editor de texto para abrir o arquivo `hue.ini`. Por exemplo:

```
sudo vim /etc/hue/conf/hue.ini
```

3. Pesquise o termo `appblacklist` e modifique a linha para o seguinte:

```
appblacklist = search
```

4. Salve as alterações e reinicie o Hue, conforme mostrado no exemplo a seguir:

```
sudo stop hue; sudo start hue
```

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

  Solução:

  - SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
  - Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

    > **Note**
    >
    > Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versões de componente 5.23.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.2.1	SDK do Amazon SageMaker Spark
emr-ddb	4.8.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.7.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.11.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.2.0	EMR S3Select Connector
emrfs	2.32.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.

Componente	Versão	Descrição
flink-client	1.7.1	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.5-amzn-3	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.5-amzn-3	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.5-amzn-3	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.5-amzn-3	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.

Componente	Versão	Descrição
hadoop-hdfs-journalnode	2.8.5-amzn-3	O serviço do HDFS para gerenciar o lançamento de arquivos do Hadoop em clusters de HA.
hadoop-httpfs-server	2.8.5-amzn-3	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.5-amzn-3	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-3	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.5-amzn-3	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.8.5-amzn-3	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.5-amzn-3	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.4.9	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.

Componente	Versão	Descrição
hbase-region-server	1.4.9	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.9	HBase cliente de linha de comando.
hbase-rest-server	1.4.9	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.9	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.4-amzn-1	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.4-amzn-1	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.3.4-amzn-1	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.4-amzn-1	O cliente da linha de comando do Hive.
hive-hbase	2.3.4-amzn-1	Cliente do Hive-hbase.
hive-metastore-server	2.3.4-amzn-1	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.

Componente	Versão	Descrição
hive-server2	2.3.4-amzn-1	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.3.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	0.9.4	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.5.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.3.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.1.0	O cliente da linha de comando do Oozie.
oozie-server	5.1.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.

Componente	Versão	Descrição
opencv	3.4.0	Biblioteca de Visão Computacional de Código Aberto.
phoenix-library	4.14.1- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.1- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.215	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.215	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística
spark-client	2.4.0	Os clientes da linha de comando do Spark.
spark-history-server	2.4.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.0	O mecanismo de execução em memória para o YARN.

Componente	Versão	Descrição
spark-yarn-slave	2.4.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	1.12.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.1	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.8.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.13	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.13	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.23.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.23.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase

Classificações	Descrição
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive- site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.pro perties do Hive.

Classificações	Descrição
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.

Classificações	Descrição
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.

Classificações	Descrição
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra .properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.

Classificações	Descrição
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.

Classificações	Descrição
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.22.0 do Amazon EMR

Versões 5.22.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hue, JupyterHub, Livy, MXNetMahout, Oozie, Phoenix, Pig, PrestoSpark, Sqoop, TensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.22.0	emr-5.21.2	emr-5.21.1	emr-5.21.0
AWS SDK for Java	1.11.510	1.11.479	1.11.479	1.11.479
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.1	1.7.0	1.7.0	1.7.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.8	1.4.8	1.4.8

	emr-5.22.0	emr-5.21.2	emr-5.21.1	emr-5.21.0
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.1	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.215	0.215	0.215
Spark	2.4.0	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1

	emr-5.22.0	emr-5.21.2	emr-5.21.1	emr-5.21.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

Notas da versão 5.22.0

As notas da versão a seguir incluem informações para a versão 5.22.0 do Amazon EMR. As alterações são referentes à versão 5.21.0.

> **Important**
>
> A partir da versão 5.22.0 do Amazon EMR, o Amazon EMR AWS usa o Signature versão 4 exclusivamente para autenticar solicitações para o Amazon S3. As versões anteriores do Amazon EMR usam o AWS Signature versão 2 em alguns casos, a menos que as notas de lançamento indiquem que o Signature versão 4 é usado exclusivamente. Para obter mais informações, consulte Autenticação de solicitações (AWS assinatura versão 4) e Solicitações de autenticação (AWS assinatura versão 2) no Guia do desenvolvedor do Amazon Simple Storage Service.

Data da versão inicial: 20 de março de 2019

Atualizações

- Flink 1.7.1
- HBase 1.4.9
- Oozie 5.1.0
- Phoenix 4.14.1
- Zeppelin 0.8.1
- Conectores e drivers:
  - Conector do DynamoDB 4.8.0
  - Conector do MariaDB 2.2.6

- Driver JDBC do Amazon Redshift, 1.2.20.1043

Novos recursos

- Modificou a configuração padrão do EBS para tipos de EC2 instância com armazenamento somente para EBS. Ao ser criado um cluster usando as versões 5.22.0 e posteriores do Amazon EMR, a quantidade padrão de armazenamento do EBS aumenta de acordo com o tamanho da instância. Além disso, dividimos o armazenamento aumentado em vários volumes, proporcionando maior performance de IOPS. Se você quiser usar uma instância de configuração de armazenamento EBS diferente, você poderá especificá-la ao criar um cluster do EMR ou adicionar nós a um cluster existente. Para obter mais informações sobre a quantidade de armazenamento e o número de volumes alocados por padrão para cada tipo de instância, consulte Armazenamento padrão do EBS para instâncias no Guia de gerenciamento do Amazon EMR.

Alterações, melhorias e problemas resolvidos

- Spark
  - Introduzida uma nova propriedade de configuração para o Spark no YARN, `spark.yarn.executor.memory0verheadFactor`. O valor dessa propriedade é um fator de escala que define o valor da sobrecarga de memória como uma porcentagem da memória do executor, com um mínimo de 384 MB. Se a sobrecarga de memória for definida explicitamente usando `spark.yarn.executor.memory0verhead`, essa propriedade não terá efeito. O valor padrão é `0.1875`, representando 18,75%. Esse padrão para o Amazon EMR deixa mais espaço nos contêineres do YARN para a sobrecarga de memória do executor do que o padrão de 10% definido internamente pelo Spark. O padrão do Amazon EMR de 18,75% mostrou empiricamente menos falhas relacionadas à memória nas avaliações comparativas do TPC-DS.
  - Backport do SPARK-26316 para melhorar a performance.
- Nas versões 5.19.0, 5.20.0 e 5.21.0 do Amazon EMR, os rótulos dos nós do YARN são armazenados em um diretório do HDFS. Em algumas situações, isso leva a atrasos na inicialização do nó central causando, em seguida, tempo limite do cluster e falha na inicialização. Desde a versão 5.22.0 do Amazon EMR, esse problema foi resolvido. Os rótulos dos nós do YARN são armazenados no disco local de cada nó do cluster, evitando dependências no HDFS.

Problemas conhecidos

- Hue (corrigido na versão 5.24.0 do Amazon EMR)

- O Hue executado no Amazon EMR não é compatível com o Solr. Desde a versão 5.20.0 do Amazon EMR, um problema de configuração incorreta faz com que o Solr seja habilitado e uma mensagem de erro inofensiva semelhante à seguinte seja exibida:

  ```
  Solr server could not be contacted properly:
  HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
  port=1978): Max retries exceeded with url: /solr/admin/info/
  system?user.name=hue&doAs=administrator&wt=json (Caused by
  NewConnectionError(': Failed to establish a new connection: [Errno 111]
  Connection refused',))
  ```

  Para evitar que a mensagem de erro do Solr seja exibida:

  1. Conecte-se à linha de comando do nó primário usando SSH.
  2. Use um editor de texto para abrir o arquivo `hue.ini`. Por exemplo:

     ```
     sudo vim /etc/hue/conf/hue.ini
     ```

  3. Pesquise o termo `appblacklist` e modifique a linha para o seguinte:

     ```
     appblacklist = search
     ```

  4. Salve as alterações e reinicie o Hue, conforme mostrado no exemplo a seguir:

     ```
     sudo stop hue; sudo start hue
     ```

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

  Solução:

  - SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
  - Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versões de componente 5.22.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.2.1	SDK do Amazon SageMaker Spark
emr-ddb	4.8.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.

Componente	Versão	Descrição
emr-goodies	2.6.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.11.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.2.0	EMR S3Select Connector
emrfs	2.31.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.7.1	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.

Componente	Versão	Descrição
hadoop-client	2.8.5-amzn-2	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.5-amzn-2	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.5-amzn-2	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.5-amzn-2	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.8.5-amzn-2	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.5-amzn-2	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.8.5-amzn-2	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.5-amzn-2	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-2	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.

Componente	Versão	Descrição
hadoop-yarn-timeline-server	2.8.5-amzn-2	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.4.9	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.9	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.9	HBase cliente de linha de comando.
hbase-rest-server	1.4.9	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.9	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.4-amzn-1	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.4-amzn-1	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.3.4-amzn-1	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.

Componente	Versão	Descrição
hive-client	2.3.4-amzn-1	O cliente da linha de comando do Hive.
hive-hbase	2.3.4-amzn-1	Cliente do Hive-hbase.
hive-metastore-server	2.3.4-amzn-1	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.4-amzn-1	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.3.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	0.9.4	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.5.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.3.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.

Componente	Versão	Descrição
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.1.0	O cliente da linha de comando do Oozie.
oozie-server	5.1.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.1- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.1- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.215	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.215	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística

Componente	Versão	Descrição
spark-client	2.4.0	Os clientes da linha de comando do Spark.
spark-history-server	2.4.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.4.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	1.12.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.1	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.8.1	O notebook baseado na web que permite um data analytics interativo.

Componente	Versão	Descrição
zookeeper-server	3.4.13	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.13	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.22.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.22.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.

Classificações	Descrição
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.

Classificações	Descrição
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.

Classificações	Descrição
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.

Classificações	Descrição
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.

Classificações	Descrição
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.

Classificações	Descrição
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.

Classificações	Descrição
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.21.2 do Amazon EMR

Versões 5.21.2 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGangliaHBase, HCatalog, HadoopHive, Hue, JupyterHub, Livy, MXNetMahout, Oozie, Phoenix, Pig, PrestoSpark, Sqoop, TensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.21.2	emr-5.21.1	emr-5.21.0	emr-5.20.1
AWS SDK for Java	1.11.479	1.11.479	1.11.479	1.11.461
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.0	1.7.0	1.7.0	1.6.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.8	1.4.8	1.4.8	1.4.8
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-

	emr-5.21.2	emr-5.21.1	emr-5.21.0	emr-5.20.1
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.215	0.215	0.214
Spark	2.4.0	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

Notas da versão 5.21.2

Esta é uma versão de patch para adicionar a autenticação AWS Signature versão 4 para solicitações ao Amazon S3. Todas as aplicações e componentes são os mesmos que os da versão anterior do Amazon EMR.

> **Important**
>
> Nesta versão de lançamento, o Amazon EMR usa o AWS Signature versão 4 exclusivamente para autenticar solicitações para o Amazon S3. Para obter mais informações, consulte Novidades.

Versões de componente 5.21.2

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.2.1	SDK do Amazon SageMaker Spark
emr-ddb	4.7.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.5.1	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.

Componente	Versão	Descrição
emr-s3-dist-cp	2.11.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.2.0	EMR S3Select Connector
emrfs	2.30.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.7.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.5-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.5-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.

Componente	Versão	Descrição
hadoop-hdfs-library	2.8.5-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.5-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.8.5-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.5-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.5-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.8.5-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.5-amzn-1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.

Componente	Versão	Descrição
hbase-hmaster	1.4.8	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.8	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.8	HBase cliente de linha de comando.
hbase-rest-server	1.4.8	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.8	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.4-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.4-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.3.4-amzn-0	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.4-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.3.4-amzn-0	Cliente do Hive-hbase.

Componente	Versão	Descrição
hive-metastore-server	2.3.4-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.4-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.3.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	0.9.4	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.5.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.3.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA

Componente	Versão	Descrição
oozie-client	5.0.0	O cliente da linha de comando do Oozie.
oozie-server	5.0.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.0- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.0- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.215	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.215	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística
spark-client	2.4.0	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	2.4.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.4.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	1.12.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.1	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.8.0	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.13	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.

Componente	Versão	Descrição
zookeeper-client	3.4.13	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.21.2

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.21.2

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.

Classificações	Descrição
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.

Classificações	Descrição
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.

Classificações	Descrição
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.

Classificações	Descrição
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.

Classificações	Descrição
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.

Classificações	Descrição
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.

Classificações	Descrição
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.21.1 do Amazon EMR

Versões 5.21.1 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.21.1	emr-5.21.0	emr-5.20.1	emr-5.20.0
AWS SDK for Java	1.11.479	1.11.479	1.11.461	1.11.461

	emr-5.21.1	emr-5.21.0	emr-5.20.1	emr-5.20.0
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.0	1.7.0	1.6.2	1.6.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.8	1.4.8	1.4.8	1.4.8
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0

	emr-5.21.1	emr-5.21.0	emr-5.20.1	emr-5.20.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.215	0.214	0.214
Spark	2.4.0	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

Notas da versão 5.21.1

As notas da versão a seguir incluem informações para a versão 5.21.1 do Amazon EMR. As alterações são referentes à versão 5.21.0.

Data da versão inicial: 18 de julho de 2019

Alterações, melhorias e problemas resolvidos

- Atualizada a AMI padrão do Amazon Linux para Amazon EMR para incluir atualizações de segurança importantes do kernel Linux, incluindo o problema de negação de serviço do TCP SACK (AWS-2019-005).

Problemas conhecidos

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

Solução:

- SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
- Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versões de componente 5.21.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`*

começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.2.1	SDK do Amazon SageMaker Spark
emr-ddb	4.7.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.5.1	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.11.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.2.0	EMR S3Select Connector
emrfs	2.30.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.7.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.5-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.5-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.5-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.5-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.8.5-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.5-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.8.5-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.

Componente	Versão	Descrição
hadoop-yarn-nodemanager	2.8.5-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hadoop-yarn-timeline-server	2.8.5-amzn-1	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.4.8	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.8	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.4.8	HBase cliente de linha de comando.
hbase-rest-server	1.4.8	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.8	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.4-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.

Componente	Versão	Descrição
hcatalog-server	2.3.4-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.3.4-amzn-0	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.4-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.3.4-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.3.4-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.4-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.3.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	0.9.4	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.5.0-incubating	Interface REST para interagir com o Apache Spark

Componente	Versão	Descrição
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.3.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.0.0	O cliente da linha de comando do Oozie.
oozie-server	5.0.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.0- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.0- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica

Componente	Versão	Descrição
presto-coordinator	0.215	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.215	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística
spark-client	2.4.0	Os clientes da linha de comando do Spark.
spark-history-server	2.4.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.4.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	1.12.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.

Componente	Versão	Descrição
tez-on-yarn	0.9.1	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.8.0	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.13	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.13	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.21.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.21.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.

Classificações	Descrição
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.

Classificações	Descrição
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.

Classificações	Descrição
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.

Classificações	Descrição
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.

Classificações	Descrição
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.

Classificações	Descrição
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.

Classificações	Descrição
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.21.0 do Amazon EMR

Versões 5.21.0 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.21.0	emr-5.20.1	emr-5.20.0	emr-5.19.1
AWS SDK for Java	1.11.479	1.11.461	1.11.461	1.11.433
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.4
Scala	2.11.12	2.11.12	2.11.12	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.0	1.6.2	1.6.2	1.6.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.8	1.4.8	1.4.8	1.4.7
HCatalog	2.3.4	2.3.4	2.3.4	2.3.3
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.3
Hudi	-	-	-	-

	emr-5.21.0	emr-5.20.1	emr-5.20.0	emr-5.19.1
Hue	4.3.0	4.3.0	4.3.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.1	1.3.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.214	0.214	0.212
Spark	2.4.0	2.4.0	2.4.0	2.3.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.11.0
Tez	0.9.1	0.9.1	0.9.1	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

Notas da versão 5.21.0

As notas da versão a seguir incluem informações para a versão 5.21.0 do Amazon EMR. As alterações são referentes à versão 5.20.0.

Data da versão inicial: 18 de fevereiro de 2019

Data da última atualização: 3 de abril de 2019

Atualizações

- Flink 1.7.0
- Presto 0.215
- AWS SDK para Java 1.11.479

Novos recursos

- (3 de abril de 2019) Com as versões 5.21.0 e posteriores do Amazon EMR, você pode substituir as configurações de cluster e especificar classificações de configuração adicionais para cada grupo de instâncias em um cluster em execução. Você faz isso usando o console do Amazon EMR, o AWS Command Line Interface (AWS CLI) ou o AWS SDK. Para obter mais informações, consulte Supplying a Configuration for an Instance Group in a Running Cluster.

Alterações, melhorias e problemas resolvidos

- Zeppelin
  - Backport do ZEPPELIN-3878.

Problemas conhecidos

- Hue (corrigido na versão 5.24.0 do Amazon EMR)
  - O Hue executado no Amazon EMR não é compatível com o Solr. Desde a versão 5.20.0 do Amazon EMR, um problema de configuração incorreta faz com que o Solr seja habilitado e uma mensagem de erro inofensiva semelhante à seguinte seja exibida:

    ```
    Solr server could not be contacted properly:
    HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
    port=1978): Max retries exceeded with url: /solr/admin/info/
    ```

```
system?user.name=hue&doAs=administrator&wt=json (Caused by
NewConnectionError(': Failed to establish a new connection: [Errno 111]
Connection refused',))
```

Para evitar que a mensagem de erro do Solr seja exibida:

1. Conecte-se à linha de comando do nó primário usando SSH.
2. Use um editor de texto para abrir o arquivo `hue.ini`. Por exemplo:

   ```
   sudo vim /etc/hue/conf/hue.ini
   ```

3. Pesquise o termo `appblacklist` e modifique a linha para o seguinte:

   ```
   appblacklist = search
   ```

4. Salve as alterações e reinicie o Hue, conforme mostrado no exemplo a seguir:

   ```
   sudo stop hue; sudo start hue
   ```

- Tez
  - Esse problema foi corrigido no Amazon EMR 5.22.0.

    Quando você se conecta à interface do usuário do Tez em http: //:8080/tez-ui *`MasterDNS`* por meio de uma conexão SSH com o nó primário do cluster, aparece o erro “Falha na operação do adaptador - o servidor Timeline (ATS) está fora de alcance. Ele está inoperante ou o CORS não está habilitado” ou as tarefas mostram, inesperadamente, N/A.

    Isso é causado pela interface do usuário do Tez ao fazer solicitações ao servidor de linha do tempo do YARN usando `localhost` em vez do nome do host do nó primário. Como solução alternativa, um script está disponível para execução como ação ou etapa de bootstrap. O script atualiza o nome do host no arquivo `configs.env` do Tez. Para obter mais informações e a localização do script, consulte Instruções de bootstrap.
- Nas versões 5.19.0, 5.20.0 e 5.21.0 do Amazon EMR, os rótulos dos nós do YARN são armazenados em um diretório do HDFS. Em algumas situações, isso leva a atrasos na inicialização do nó central causando, em seguida, tempo limite do cluster e falha na inicialização. Desde a versão 5.22.0 do Amazon EMR, esse problema foi resolvido. Os rótulos dos nós do YARN são armazenados no disco local de cada nó do cluster, evitando dependências no HDFS.
- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

Solução:

- SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
- Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versões de componente 5.21.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`*

começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.2.1	SDK do Amazon SageMaker Spark
emr-ddb	4.7.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.5.1	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.11.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.2.0	EMR S3Select Connector
emrfs	2.30.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.7.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.5-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.5-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.5-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.5-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.8.5-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.5-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.8.5-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.

Componente	Versão	Descrição
hadoop-yarn-nodemanager	2.8.5-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hadoop-yarn-timeline-server	2.8.5-amzn-1	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.4.8	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.8	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.4.8	HBase cliente de linha de comando.
hbase-rest-server	1.4.8	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.8	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.4-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.

Componente	Versão	Descrição
hcatalog-server	2.3.4-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.3.4-amzn-0	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.4-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.3.4-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.3.4-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.4-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.3.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	0.9.4	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.5.0-incubating	Interface REST para interagir com o Apache Spark

Componente	Versão	Descrição
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.3.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.0.0	O cliente da linha de comando do Oozie.
oozie-server	5.0.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.0- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.0- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica

Componente	Versão	Descrição
presto-coordinator	0.215	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.215	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística
spark-client	2.4.0	Os clientes da linha de comando do Spark.
spark-history-server	2.4.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.4.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	1.12.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.

Componente	Versão	Descrição
tez-on-yarn	0.9.1	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.8.0	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.13	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.13	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.21.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.21.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.

Classificações	Descrição
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.

Classificações	Descrição
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.

Classificações	Descrição
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.

Classificações	Descrição
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.

Classificações	Descrição
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.

Classificações	Descrição
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.

Classificações	Descrição
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.20.1 do Amazon EMR

Versões 5.20.1 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.20.1	emr-5.20.0	emr-5.19.1	emr-5.19.0
AWS SDK for Java	1.11.461	1.11.461	1.11.433	1.11.433
Python	2.7, 3.6	2.7, 3.6	2.7, 3.4	2.7, 3.4
Scala	2.11.12	2.11.12	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.6.2	1.6.2	1.6.1	1.6.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.8	1.4.8	1.4.7	1.4.7
HCatalog	2.3.4	2.3.4	2.3.3	2.3.3
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.3	2.3.3
Hudi	-	-	-	-

	emr-5.20.1	emr-5.20.0	emr-5.19.1	emr-5.19.0
Hue	4.3.0	4.3.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.0	1.3.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.214	0.214	0.212	0.212
Spark	2.4.0	2.4.0	2.3.2	2.3.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.11.0	1.11.0
Tez	0.9.1	0.9.1	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

Notas da versão 5.20.1

Esta é uma versão de patch para adicionar a autenticação AWS Signature versão 4 para solicitações ao Amazon S3. Todas as aplicações e componentes são os mesmos que os da versão anterior do Amazon EMR.

> **Important**
>
> Nesta versão de lançamento, o Amazon EMR usa o AWS Signature versão 4 exclusivamente para autenticar solicitações para o Amazon S3. Para obter mais informações, consulte Novidades.

Versões de componente 5.20.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.2.1	SDK do Amazon SageMaker Spark
emr-ddb	4.7.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.5.1	Bibliotecas convenientes para o ecossistema do Hadoop.

Componente	Versão	Descrição
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.10.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.2.0	EMR S3Select Connector
emrfs	2.29.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.6.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.5-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.

Componente	Versão	Descrição
hadoop-hdfs-datanode	2.8.5-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.5-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.5-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.8.5-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.5-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.5-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.8.5-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.5-amzn-1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.

Componente	Versão	Descrição
hbase-hmaster	1.4.8	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.8	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.4.8	HBase cliente de linha de comando.
hbase-rest-server	1.4.8	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.8	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.4-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.4-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.3.4-amzn-0	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.4-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.3.4-amzn-0	Cliente do Hive-hbase.

Componente	Versão	Descrição
hive-metastore-server	2.3.4-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.4-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.3.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	0.9.4	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.5.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.3.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA

Componente	Versão	Descrição
oozie-client	5.0.0	O cliente da linha de comando do Oozie.
oozie-server	5.0.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.0- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.0- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.214	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.214	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística
spark-client	2.4.0	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	2.4.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.4.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	1.12.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.1	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.8.0	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.13	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.

Componente	Versão	Descrição
zookeeper-client	3.4.13	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.20.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.20.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.

Classificações	Descrição
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.

Classificações	Descrição
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.

Classificações	Descrição
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.

Classificações	Descrição
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.

Classificações	Descrição
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.

Classificações	Descrição
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.

Classificações	Descrição
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.20.0 do Amazon EMR

Versões 5.20.0 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.20.0	emr-5.19.1	emr-5.19.0	emr-5.18.1
AWS SDK for Java	1.11.461	1.11.433	1.11.433	1.11.393

	emr-5.20.0	emr-5.19.1	emr-5.19.0	emr-5.18.1
Python	2.7, 3.6	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.12	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.6.2	1.6.1	1.6.1	1.6.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.8	1.4.7	1.4.7	1.4.7
HCatalog	2.3.4	2.3.3	2.3.3	2.3.3
Hadoop	2.8.5	2.8.5	2.8.5	2.8.4
Hive	2.3.4	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.3.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.0	1.3.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0

	emr-5.20.0	emr-5.19.1	emr-5.19.0	emr-5.18.1
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.214	0.212	0.212	0.210
Spark	2.4.0	2.3.2	2.3.2	2.3.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.11.0	1.11.0	1.9.0
Tez	0.9.1	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.12

Notas da versão 5.20.0

As notas da versão a seguir incluem informações para a versão 5.20.0 do Amazon EMR. As alterações são referentes à versão 5.19.0.

Data da versão inicial: 18 de dezembro de 2018

Data da última atualização: 22 de janeiro de 2019

Atualizações

- Flink 1.6.2
- HBase 1.4.8
- Hive 2.3.4
- Hue 4.3.0
- MXNet 1.3.1
- Presto 0.214

- Spark 2.4.0
- TensorFlow 1.12.0
- Tez 0.9.1
- AWS SDK para Java 1.11.461

Novos recursos

- (22 de janeiro de 2019) O Kerberos no Amazon EMR foi aprimorado para oferecer suporte à autenticação de entidades principais de um KDC externo. Isso centraliza o gerenciamento de principais porque vários clusters podem compartilhar um único KDC externo. Além disso, o KDC externo pode ter uma relação de confiança entre realm com um domínio do Active Directory. Isso permite que todos os clusters autentiquem principais do Active Directory. Para obter mais informações, consulte Usar autenticação Kerberos no Guia de gerenciamento do Amazon EMR.

Alterações, melhorias e problemas resolvidos

- AMI padrão do Amazon Linux para Amazon EMR
  - O pacote Python3 foi atualizado do python 3.4 para 3.6.
- O confirmador otimizado para EMRFS S3
  - O confirmador otimizado para EMRFS S3 agora está habilitado por padrão, o que melhora a performance de gravação. Para obter mais informações, consulte Usar o confirmador otimizado para EMRFS S3.
- Hive
  - Backport do HIVE-16686.
- Glue com Spark e Hive
  - No EMR 5.20.0 ou posterior, a remoção paralela de partições é ativada automaticamente para Spark e Hive quando o AWS Glue Data Catalog é usado como metastore. Essa alteração reduz significativamente o tempo de planejamento de consultas ao executar várias solicitações em paralelo para recuperar partições. O número total de segmentos que podem ser executados simultaneamente varia entre 1 e 10. O valor padrão é 5, que é uma configuração recomendada. Você pode alterá-lo especificando a propriedade `aws.glue.partition.num.segments` na classificação de configuração `hive-site`. Se ocorrer controle de utilização, você poderá desativar o atributo alterando o valor para 1. Para obter mais informações, consulte a Estrutura de segmentos do AWS Glue.

Problemas conhecidos

- Hue (corrigido na versão 5.24.0 do Amazon EMR)
  - O Hue executado no Amazon EMR não é compatível com o Solr. Desde a versão 5.20.0 do Amazon EMR, um problema de configuração incorreta faz com que o Solr seja habilitado e uma mensagem de erro inofensiva semelhante à seguinte seja exibida:

    ```
    Solr server could not be contacted properly:
    HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
    port=1978): Max retries exceeded with url: /solr/admin/info/
    system?user.name=hue&doAs=administrator&wt=json (Caused by
    NewConnectionError(': Failed to establish a new connection: [Errno 111]
    Connection refused',))
    ```

    Para evitar que a mensagem de erro do Solr seja exibida:

    1. Conecte-se à linha de comando do nó primário usando SSH.
    2. Use um editor de texto para abrir o arquivo `hue.ini`. Por exemplo:

       ```
       sudo vim /etc/hue/conf/hue.ini
       ```

    3. Pesquise o termo `appblacklist` e modifique a linha para o seguinte:

       ```
       appblacklist = search
       ```

    4. Salve as alterações e reinicie o Hue, conforme mostrado no exemplo a seguir:

       ```
       sudo stop hue; sudo start hue
       ```

- Tez
  - Esse problema foi corrigido no Amazon EMR 5.22.0.

    Quando você se conecta à interface do usuário do Tez em http: //:8080/tez-ui *MasterDNS* por meio de uma conexão SSH com o nó primário do cluster, aparece o erro “Falha na operação do adaptador - o servidor Timeline (ATS) está fora de alcance. Ele está inoperante ou o CORS não está habilitado” ou as tarefas mostram, inesperadamente, N/A.

    Isso é causado pela interface do usuário do Tez ao fazer solicitações ao servidor de linha do tempo do YARN usando `localhost` em vez do nome do host do nó primário. Como solução alternativa, um script está disponível para execução como ação ou etapa de bootstrap. O script

atualiza o nome do host no arquivo `configs.env` do Tez. Para obter mais informações e a localização do script, consulte Instruções de bootstrap.

- Nas versões 5.19.0, 5.20.0 e 5.21.0 do Amazon EMR, os rótulos dos nós do YARN são armazenados em um diretório do HDFS. Em algumas situações, isso leva a atrasos na inicialização do nó central causando, em seguida, tempo limite do cluster e falha na inicialização. Desde a versão 5.22.0 do Amazon EMR, esse problema foi resolvido. Os rótulos dos nós do YARN são armazenados no disco local de cada nó do cluster, evitando dependências no HDFS.
- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

  Solução:

  - SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
  - Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versões de componente 5.20.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.2.1	SDK do Amazon SageMaker Spark
emr-ddb	4.7.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.5.1	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.10.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.2.0	EMR S3Select Connector
emrfs	2.29.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.

Componente	Versão	Descrição
flink-client	1.6.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.5-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.5-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.5-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.5-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.8.5-amzn-1	O endpoint de HTTP para as operações do HDFS.

Componente	Versão	Descrição
hadoop-kms-server	2.8.5-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.5-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.8.5-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.5-amzn-1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.4.8	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.8	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.4.8	HBase cliente de linha de comando.

Componente	Versão	Descrição
hbase-rest-server	1.4.8	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.8	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.4-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.4-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.3.4-amzn-0	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.4-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.3.4-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.3.4-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.4-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.

Componente	Versão	Descrição
hue-server	4.3.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	0.9.4	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.5.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.3.1	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.0.0	O cliente da linha de comando do Oozie.
oozie-server	5.0.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computaci onal de Código Aberto.

Componente	Versão	Descrição
phoenix-library	4.14.0- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.0- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.214	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.214	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística
spark-client	2.4.0	Os clientes da linha de comando do Spark.
spark-history-server	2.4.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.4.0	O mecanismo de execução em memória para o YARN.

Componente	Versão	Descrição
spark-yarn-slave	2.4.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	1.12.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.9.1	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.8.0	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.13	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.13	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.20.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.20.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase

Classificações	Descrição
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive- site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.pro perties do Hive.

Classificações	Descrição
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.

Classificações	Descrição
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.

Classificações	Descrição
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra .properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.

Classificações	Descrição
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.

Classificações	Descrição
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.19.1 do Amazon EMR

Versões 5.19.1 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.19.1	emr-5.19.0	emr-5.18.1	emr-5.18.0
AWS SDK for Java	1.11.433	1.11.433	1.11.393	1.11.393
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.6.1	1.6.1	1.6.0	1.6.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.7	1.4.7	1.4.7	1.4.7

	emr-5.19.1	emr-5.19.0	emr-5.18.1	emr-5.18.0
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.5	2.8.5	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.0	1.3.0	1.2.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.212	0.212	0.210	0.210
Spark	2.3.2	2.3.2	2.3.2	2.3.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.11.0	1.11.0	1.9.0	1.9.0
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.19.1	emr-5.19.0	emr-5.18.1	emr-5.18.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.12	3.4.12

Notas da versão 5.19.1

Esta é uma versão de patch para adicionar a autenticação AWS Signature versão 4 para solicitações ao Amazon S3. Todas as aplicações e componentes são os mesmos que os da versão anterior do Amazon EMR.

> **Important**
>
> Nesta versão de lançamento, o Amazon EMR usa o AWS Signature versão 4 exclusivamente para autenticar solicitações para o Amazon S3. Para obter mais informações, consulte Novidades.

Versões de componente 5.19.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.2.0	SDK do Amazon SageMaker Spark
emr-ddb	4.7.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.5.1	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.10.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.1.0	EMR S3Select Connector
emrfs	2.28.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.6.1	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.5-amzn-0	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.5-amzn-0	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.5-amzn-0	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.5-amzn-0	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.8.5-amzn-0	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.5-amzn-0	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.8.5-amzn-0	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.5-amzn-0	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.8.5-amzn-0	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.5-amzn-0	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.4.7	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.7	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.4.7	HBase cliente de linha de comando.
hbase-rest-server	1.4.7	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.7	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.3-amzn-2	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.3-amzn-2	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	2.3.3-amzn-2	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.3-amzn-2	O cliente da linha de comando do Hive.
hive-hbase	2.3.3-amzn-2	Cliente do Hive-hbase.
hive-metastore-server	2.3.3-amzn-2	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.3-amzn-2	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.2.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	0.9.4	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.5.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.

Componente	Versão	Descrição
mxnet	1.3.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.0.0	O cliente da linha de comando do Oozie.
oozie-server	5.0.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.0- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.0- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.212	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.212	O serviço que executa partes de uma consulta.

Componente	Versão	Descrição
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística
spark-client	2.3.2	Os clientes da linha de comando do Spark.
spark-history-server	2.3.2	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.3.2	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.3.2	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	1.11.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.

Componente	Versão	Descrição
zeppelin-server	0.8.0	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.13	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.13	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.19.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.19.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.

Classificações	Descrição
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.

Classificações	Descrição
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.

Classificações	Descrição
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.

Classificações	Descrição
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.

Classificações	Descrição
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.

Classificações	Descrição
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.

Classificações	Descrição
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.19.0 do Amazon EMR

Versões 5.19.0 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.19.0	emr-5.18.1	emr-5.18.0	emr-5.17.2
AWS SDK for Java	1.11.433	1.11.393	1.11.393	1.11.336
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.6.1	1.6.0	1.6.0	1.5.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.7	1.4.7	1.4.7	1.4.6
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.5	2.8.4	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-

	emr-5.19.0	emr-5.18.1	emr-5.18.0	emr-5.17.2
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.0	1.2.0	1.2.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.212	0.210	0.210	0.206
Spark	2.3.2	2.3.2	2.3.2	2.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.11.0	1.9.0	1.9.0	1.9.0
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.7.3
ZooKeeper	3.4.13	3.4.12	3.4.12	3.4.12

Notas da versão 5.19.0

As notas da versão a seguir incluem informações para a versão 5.19.0 do Amazon EMR. As alterações são referentes à versão 5.18.0.

Data da versão inicial: 7 de novembro de 2018

Data da última atualização: 19 de novembro de 2018

Atualizações

- Hadoop 2.8.5
- Flink 1.6.1
- JupyterHub 0.9.4
- MXNet 1.3.0
- Presto 0.212
- TensorFlow 1.11.0
- Zookeeper 3.4.13
- AWS SDK para Java 1.11.433

Novos recursos

- (19 de novembro de 2018) Os Cadernos do EMR constituem um ambiente gerenciado baseado no caderno Jupyter. Ele suporta os kernels mágicos do Spark para Spark PySpark SQL, Spark R e Scala. Os Cadernos do EMR podem ser usados com clusters criados usando as versões 5.18.0 e posteriores do Amazon EMR. Para obter mais informações, consulte Usar Cadernos do EMR no Guia de gerenciamento do Amazon EMR.
- O confirmador otimizado para EMRFS S3 está disponível ao serem gravados arquivos Parquet usando Spark e EMRFS. Esse confirmador melhora a performance de gravação. Para obter mais informações, consulte Usar o confirmador otimizado para EMRFS S3.

Alterações, melhorias e problemas resolvidos

- YARN
  - Modificada a lógica que limita o processo mestre da aplicação à execução nos nós centrais. Essa funcionalidade agora usa o atributo e as propriedades de rótulos de nós do YARN nas classificações de configuração `yarn-site` e `capacity-scheduler`. Para obter mais informações, consulte https://docs.aws.amazon.com/emr/latest/ManagementGuide/emr-plan-instances-guidelines.html#emr-plan-spot-YARN.
- AMI padrão do Amazon Linux para Amazon EMR

- `ruby18`, `php56`, e `gcc48` não são mais instalados por padrão. Eles podem ser instalados, se desejado, usando `yum`.
- A gem do ruby aws-sdk não é mais instalada por padrão. Ela pode ser instalada usando `gem install aws-sdk`, se desejado. Componentes específicos também podem ser instalados. Por exemplo, `.gem install aws-sdk-s3`

Problemas conhecidos

- Cadernos do EMR: em algumas circunstâncias, com vários editores de cadernos abertos, o editor de cadernos pode parecer incapaz de se conectar ao cluster. Se isso acontecer, limpe os cookies do navegador e reabra os editores de cadernos.
- CloudWatch ContainerPending Escalabilidade métrica e automática — (corrigida na versão 5.20.0) O Amazon EMR pode emitir um valor negativo para. `ContainerPending` Se `ContainerPending` for usado em uma regra de escalabilidade automática, a escalabilidade automática não se comportará conforme esperado. Evite usar `ContainerPending` com escalabilidade automática.
- Nas versões 5.19.0, 5.20.0 e 5.21.0 do Amazon EMR, os rótulos dos nós do YARN são armazenados em um diretório do HDFS. Em algumas situações, isso leva a atrasos na inicialização do nó central causando, em seguida, tempo limite do cluster e falha na inicialização. Desde a versão 5.22.0 do Amazon EMR, esse problema foi resolvido. Os rótulos dos nós do YARN são armazenados no disco local de cada nó do cluster, evitando dependências no HDFS.

Versões de componente 5.19.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.2.0	SDK do Amazon SageMaker Spark
emr-ddb	4.7.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.5.1	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.10.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.1.0	EMR S3Select Connector
emrfs	2.28.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.6.1	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.5-amzn-0	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.5-amzn-0	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.5-amzn-0	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.5-amzn-0	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.8.5-amzn-0	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.5-amzn-0	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.8.5-amzn-0	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.5-amzn-0	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.8.5-amzn-0	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.5-amzn-0	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.4.7	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.7	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.4.7	HBase cliente de linha de comando.
hbase-rest-server	1.4.7	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.7	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.3-amzn-2	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.3-amzn-2	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	2.3.3-amzn-2	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.3-amzn-2	O cliente da linha de comando do Hive.
hive-hbase	2.3.3-amzn-2	Cliente do Hive-hbase.
hive-metastore-server	2.3.3-amzn-2	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.3-amzn-2	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.2.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	0.9.4	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.5.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.

Componente	Versão	Descrição
mxnet	1.3.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.0.0	O cliente da linha de comando do Oozie.
oozie-server	5.0.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.0- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.0- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.212	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.212	O serviço que executa partes de uma consulta.

Componente	Versão	Descrição
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística
spark-client	2.3.2	Os clientes da linha de comando do Spark.
spark-history-server	2.3.2	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.3.2	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.3.2	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	1.11.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.

Componente	Versão	Descrição
zeppelin-server	0.8.0	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.13	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.13	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.19.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.19.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.

Classificações	Descrição
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.

Classificações	Descrição
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.

Classificações	Descrição
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.

Classificações	Descrição
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.

Classificações	Descrição
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-memory	Alterar os valores no arquivo memory.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.

Classificações	Descrição
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
presto-connector-tpcds	Alterar os valores no arquivo tpcds.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.

Classificações	Descrição
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.18.1 do Amazon EMR

Versões 5.18.1 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.18.1	emr-5.18.0	emr-5.17.2	emr-5.17.1
AWS SDK for Java	1.11.393	1.11.393	1.11.336	1.11.336
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.6.0	1.6.0	1.5.2	1.5.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.7	1.4.7	1.4.6	1.4.6
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-

	emr-5.18.1	emr-5.18.0	emr-5.17.2	emr-5.17.1
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.2.0	1.2.0	1.2.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.210	0.210	0.206	0.206
Spark	2.3.2	2.3.2	2.3.1	2.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.9.0	1.9.0	1.9.0	1.9.0
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

Notas da versão 5.18.1

Esta é uma versão de patch para adicionar a autenticação AWS Signature versão 4 para solicitações ao Amazon S3. Todas as aplicações e componentes são os mesmos que os da versão anterior do Amazon EMR.

> **Important**
>
> Nesta versão de lançamento, o Amazon EMR usa o AWS Signature versão 4 exclusivamente para autenticar solicitações para o Amazon S3. Para obter mais informações, consulte Novidades.

Versões de componente 5.18.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.1.3	SDK do Amazon SageMaker Spark
emr-ddb	4.6.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.5.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.

Componente	Versão	Descrição
emr-s3-dist-cp	2.10.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.1.0	EMR S3Select Connector
emrfs	2.27.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.6.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.4-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.4-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.

Componente	Versão	Descrição
hadoop-hdfs-library	2.8.4-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.4-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.8.4-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.4-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.8.4-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.4-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.8.4-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.4-amzn-1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.

Componente	Versão	Descrição
hbase-hmaster	1.4.7	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.7	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.4.7	HBase cliente de linha de comando.
hbase-rest-server	1.4.7	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.7	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.3-amzn-2	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.3-amzn-2	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.3.3-amzn-2	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.3-amzn-2	O cliente da linha de comando do Hive.
hive-hbase	2.3.3-amzn-2	Cliente do Hive-hbase.

Componente	Versão	Descrição
hive-metastore-server	2.3.3-amzn-2	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.3-amzn-2	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.2.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	0.8.1	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.5.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.2.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA

Componente	Versão	Descrição
oozie-client	5.0.0	O cliente da linha de comando do Oozie.
oozie-server	5.0.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.0- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.0- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.210	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.210	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística
spark-client	2.3.2	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	2.3.2	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.3.2	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.3.2	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	1.9.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.8.0	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.12	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.

Componente	Versão	Descrição
zookeeper-client	3.4.12	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.18.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.18.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.

Classificações	Descrição
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.

Classificações	Descrição
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.

Classificações	Descrição
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.

Classificações	Descrição
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.

Classificações	Descrição
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.

Classificações	Descrição
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.

Classificações	Descrição
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.18.0 do Amazon EMR

Versões 5.18.0 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.18.0	emr-5.17.2	emr-5.17.1	emr-5.17.0
AWS SDK for Java	1.11.393	1.11.336	1.11.336	1.11.336
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8

	emr-5.18.0	emr-5.17.2	emr-5.17.1	emr-5.17.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.6.0	1.5.2	1.5.2	1.5.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.7	1.4.6	1.4.6	1.4.6
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.2.0	1.2.0	1.2.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0

	emr-5.18.0	emr-5.17.2	emr-5.17.1	emr-5.17.0
Presto	0.210	0.206	0.206	0.206
Spark	2.3.2	2.3.1	2.3.1	2.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.9.0	1.9.0	1.9.0	1.9.0
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

Notas da versão 5.18.0

As notas da versão a seguir incluem informações para a versão 5.18.0 do Amazon EMR. As alterações são referentes à versão 5.17.0.

Data da versão inicial: 24 de outubro de 2018

Atualizações

- Flink 1.6.0
- HBase 1.4.7
- Presto 0.210
- Spark 2.3.2
- Zeppelin 0.8.0

Novos recursos

- Desde a versão 5.18.0 do Amazon EMR, você pode usar o repositório de artefatos do Amazon EMR para criar o código de trabalho em comparação com as versões exatas de bibliotecas

e dependências disponíveis com versões específicas do Amazon EMR. Para obter mais informações, consulte Verificar dependências usando o repositório de artefatos do Amazon EMR.

Alterações, melhorias e problemas resolvidos

- Hive
  - Adicionado suporte para o S3 Select. Para obter mais informações, consulte Usar o S3 Select com o Hive para melhorar a performance.
- Presto
  - Adicionado suporte para o S3 Select Pushdown. Para obter mais informações, consulte Usar S3 Select Pushdown com o Presto para melhorar a performance.
- Spark
  - A configuração log4j padrão do Spark foi alterada para lançar logs de contêineres por hora para trabalhos de streaming do Spark. Isso ajuda a evitar a exclusão de lo de trabalhos de streaming do Spark de execução prolongada.

Versões de componente 5.18.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.1.3	SDK do Amazon SageMaker Spark

Componente	Versão	Descrição
emr-ddb	4.6.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.5.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.10.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.1.0	EMR S3Select Connector
emrfs	2.27.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.6.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.4-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.4-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.4-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.4-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.8.4-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.4-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.8.4-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.4-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.8.4-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.4-amzn-1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.4.7	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.7	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.7	HBase cliente de linha de comando.
hbase-rest-server	1.4.7	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.7	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.3-amzn-2	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.3-amzn-2	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	2.3.3-amzn-2	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.3-amzn-2	O cliente da linha de comando do Hive.
hive-hbase	2.3.3-amzn-2	Cliente do Hive-hbase.
hive-metastore-server	2.3.3-amzn-2	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.3-amzn-2	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.2.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	0.8.1	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.5.0-incubating	Interface REST para interagir com o Apache Spark
nginx	1.12.1	nginx [mecanismo x] é um servidor de proxy reverso e HTTP
mahout-client	0.13.0	A biblioteca de Machine Learning.

Componente	Versão	Descrição
mxnet	1.2.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.0.0	O cliente da linha de comando do Oozie.
oozie-server	5.0.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.0- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.0- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.210	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.210	O serviço que executa partes de uma consulta.

Componente	Versão	Descrição
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística
spark-client	2.3.2	Os clientes da linha de comando do Spark.
spark-history-server	2.3.2	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.3.2	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.3.2	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	1.9.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.

Componente	Versão	Descrição
zeppelin-server	0.8.0	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.12	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.12	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.18.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.18.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.

Classificações	Descrição
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.

Classificações	Descrição
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.

Classificações	Descrição
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.

Classificações	Descrição
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.

Classificações	Descrição
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.

Classificações	Descrição
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.

Classificações	Descrição
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.17.2 do Amazon EMR

Versões 5.17.2 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.17.2	emr-5.17.1	emr-5.17.0	emr-5.16.1
AWS SDK for Java	1.11.336	1.11.336	1.11.336	1.11.336
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.5.2	1.5.2	1.5.2	1.5.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.6	1.4.6	1.4.6	1.4.4
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.2.0	1.2.0	1.2.0	1.2.0

	emr-5.17.2	emr-5.17.1	emr-5.17.0	emr-5.16.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.206	0.206	0.206	0.203
Spark	2.3.1	2.3.1	2.3.1	2.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.9.0	1.9.0	1.9.0	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

Notas da versão 5.17.2

Esta é uma versão de patch para adicionar a autenticação AWS Signature versão 4 para solicitações ao Amazon S3. Todas as aplicações e componentes são os mesmos que os da versão anterior do Amazon EMR.

> **Important**
>
> Nesta versão de lançamento, o Amazon EMR usa o AWS Signature versão 4 exclusivamente para autenticar solicitações para o Amazon S3. Para obter mais informações, consulte Novidades.

Versões de componente 5.17.2

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.1.3	SDK do Amazon SageMaker Spark
emr-ddb	4.6.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.5.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.10.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.0.0	EMR S3Select Connector
emrfs	2.26.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.

Componente	Versão	Descrição
flink-client	1.5.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.4-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.4-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.4-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.4-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.8.4-amzn-1	O endpoint de HTTP para as operações do HDFS.

Componente	Versão	Descrição
hadoop-kms-server	2.8.4-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.8.4-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.4-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.8.4-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.4-amzn-1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.4.6	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.6	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.6	HBase cliente de linha de comando.

Componente	Versão	Descrição
hbase-rest-server	1.4.6	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.6	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.3-amzn-1	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.3-amzn-1	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.3.3-amzn-1	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.3-amzn-1	O cliente da linha de comando do Hive.
hive-hbase	2.3.3-amzn-1	Cliente do Hive-hbase.
hive-metastore-server	2.3.3-amzn-1	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.3-amzn-1	O serviço que aceita as consultas do Hive como solicitações da web.

Componente	Versão	Descrição
hue-server	4.2.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	0.8.1	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.5.0-incubating	Interface REST para interagir com o Apache Spark
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.2.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.0.0	O cliente da linha de comando do Oozie.
oozie-server	5.0.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.0- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente

Componente	Versão	Descrição
phoenix-query-server	4.14.0- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.206	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.206	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística
spark-client	2.3.1	Os clientes da linha de comando do Spark.
spark-history-server	2.3.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.3.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.3.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.

Componente	Versão	Descrição
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	1.9.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.3	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.12	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.12	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.17.2

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.17.2

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase

Classificações	Descrição
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive- site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.pro perties do Hive.

Classificações	Descrição
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.

Classificações	Descrição
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.

Classificações	Descrição
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra .properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.

Classificações	Descrição
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.

Classificações	Descrição
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.17.1 do Amazon EMR

Versões 5.17.1 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.17.1	emr-5.17.0	emr-5.16.1	emr-5.16.0
AWS SDK for Java	1.11.336	1.11.336	1.11.336	1.11.336
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.5.2	1.5.2	1.5.0	1.5.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.6	1.4.6	1.4.4	1.4.4

	emr-5.17.1	emr-5.17.0	emr-5.16.1	emr-5.16.0
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.2.0	1.2.0	1.2.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.206	0.206	0.203	0.203
Spark	2.3.1	2.3.1	2.3.1	2.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.9.0	1.9.0	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.17.1	emr-5.17.0	emr-5.16.1	emr-5.16.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

Notas da versão 5.17.1

As notas da versão a seguir incluem informações para a versão 5.17.1 do Amazon EMR. As alterações são referentes à versão 5.17.0.

Data da versão inicial: 18 de julho de 2019

Alterações, melhorias e problemas resolvidos

- Atualizada a AMI padrão do Amazon Linux para Amazon EMR para incluir atualizações de segurança importantes do kernel Linux, incluindo o problema de negação de serviço do TCP SACK (AWS-2019-005).

Versões de componente 5.17.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.1.3	SDK do Amazon SageMaker Spark
emr-ddb	4.6.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.5.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.10.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.0.0	EMR S3Select Connector
emrfs	2.26.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.5.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.4-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.4-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.4-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.4-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.8.4-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.4-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.8.4-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.4-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.8.4-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.4-amzn-1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.4.6	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.6	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.6	HBase cliente de linha de comando.
hbase-rest-server	1.4.6	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.6	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.3-amzn-1	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.3-amzn-1	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	2.3.3-amzn-1	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.3-amzn-1	O cliente da linha de comando do Hive.
hive-hbase	2.3.3-amzn-1	Cliente do Hive-hbase.
hive-metastore-server	2.3.3-amzn-1	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.3-amzn-1	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.2.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	0.8.1	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.5.0-incubating	Interface REST para interagir com o Apache Spark
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.2.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.

Componente	Versão	Descrição
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.0.0	O cliente da linha de comando do Oozie.
oozie-server	5.0.0	O serviço que aceita solicitações de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computacional de Código Aberto.
phoenix-library	4.14.0- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.0- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.206	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.206	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística

Componente	Versão	Descrição
spark-client	2.3.1	Os clientes da linha de comando do Spark.
spark-history-server	2.3.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.3.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.3.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	1.9.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.3	O notebook baseado na web que permite um data analytics interativo.

Componente	Versão	Descrição
zookeeper-server	3.4.12	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.12	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.17.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.17.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.

Classificações	Descrição
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.

Classificações	Descrição
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.

Classificações	Descrição
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.

Classificações	Descrição
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.

Classificações	Descrição
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.

Classificações	Descrição
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.

Classificações	Descrição
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.17.0 do Amazon EMR

Versões 5.17.0 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.17.0	emr-5.16.1	emr-5.16.0	emr-5.15.1
AWS SDK for Java	1.11.336	1.11.336	1.11.336	1.11.33
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.5.2	1.5.0	1.5.0	1.4.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.6	1.4.4	1.4.4	1.4.4
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.4	2.8.3
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.4.0
MXNet	1.2.0	1.2.0	1.2.0	1.1.0

	emr-5.17.0	emr-5.16.1	emr-5.16.0	emr-5.15.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.206	0.203	0.203	0.194
Spark	2.3.1	2.3.1	2.3.1	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.9.0	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

Notas da versão 5.17.0

As notas da versão a seguir incluem informações para a versão 5.17.0 do Amazon EMR. As alterações são referentes à versão 5.16.0.

Data da versão inicial: 30 de agosto de 2018

Atualizações

- Flink 1.5.2
- HBase 1.4.6
- Presto 0.206

Novos recursos

- Adicionado suporte para Tensorflow. Para obter mais informações, consulte TensorFlow.

Alterações, melhorias e problemas resolvidos

- JupyterHub
  - Adicionado suporte para a persistência de cadernos no Amazon S3. Para obter mais informações, consulte Configurar a persistência de cadernos no Amazon S3.
- Spark
  - Adicionado suporte para o S3 Select. Para obter mais informações, consulte Usar o S3 Select com Spark para melhorar a performance das consultas.
- Resolvidos os problemas com as métricas do Cloudwatch e o atributo de escalabilidade automática nas versões 5.14.0, 5.15.0 ou 5.16.0 do Amazon EMR.

Problemas conhecidos

- Quando você cria um cluster kerberizado com o Livy instalado, o Livy apresenta falha com um erro em que a autenticação simples não está habilitada. A reinicialização do servidor do Livy resolve o problema. Como solução alternativa, adicione uma etapa durante a criação do cluster que execute `sudo restart livy-server` no nó primário.
- Se você usar uma AMI do Amazon Linux personalizada com base em uma AMI do Amazon Linux com data de criação 11/8/2018, o servidor Oozie falhará ao iniciar. Se você usar o Oozie, crie uma AMI personalizada com base em um ID de AMI do Amazon Linux com uma data de criação diferente. Você pode usar o AWS CLI comando a seguir para retornar uma lista de imagens IDs para todos os HVM Amazon Linux AMIs com uma versão 2018.03, junto com a data de lançamento, para que você possa escolher uma Amazon Linux AMI apropriada como sua base. MyRegion Substitua pelo seu identificador de região, como us-west-2.

```
aws ec2 --region MyRegion describe-images --owner amazon --query 'Images[?
Name!=`null`]|[?starts_with(Name, `amzn-ami-hvm-2018.03`) == `true`].
[CreationDate,ImageId,Name]' --output text | sort -rk1
```

Versões de componente 5.17.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.1.3	SDK do Amazon SageMaker Spark
emr-ddb	4.6.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.5.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.10.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emr-s3-select	1.0.0	EMR S3Select Connector
emrfs	2.26.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.

Componente	Versão	Descrição
flink-client	1.5.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.4-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.4-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.4-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.4-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.8.4-amzn-1	O endpoint de HTTP para as operações do HDFS.

Componente	Versão	Descrição
hadoop-kms-server	2.8.4-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.8.4-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.4-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.8.4-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.4-amzn-1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.4.6	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.6	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.6	HBase cliente de linha de comando.

Componente	Versão	Descrição
hbase-rest-server	1.4.6	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.6	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.3-amzn-1	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.3-amzn-1	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.3.3-amzn-1	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.3-amzn-1	O cliente da linha de comando do Hive.
hive-hbase	2.3.3-amzn-1	Cliente do Hive-hbase.
hive-metastore-server	2.3.3-amzn-1	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.3-amzn-1	O serviço que aceita as consultas do Hive como solicitações da web.

Componente	Versão	Descrição
hue-server	4.2.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	0.8.1	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.5.0-incubating	Interface REST para interagir com o Apache Spark
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.2.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.0.0	O cliente da linha de comando do Oozie.
oozie-server	5.0.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.0- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente

Componente	Versão	Descrição
phoenix-query-server	4.14.0- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.206	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.206	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística
spark-client	2.3.1	Os clientes da linha de comando do Spark.
spark-history-server	2.3.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.3.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.3.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.

Componente	Versão	Descrição
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tensorflow	1.9.0	TensorFlow biblioteca de software de código aberto para computação numérica de alto desempenho.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.3	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.12	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.12	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.17.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.17.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase

Classificações	Descrição
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive- site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.pro perties do Hive.

Classificações	Descrição
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.

Classificações	Descrição
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-s3-conf	Configurar a persistência do notebook Jupyter S3.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.

Classificações	Descrição
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra .properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.

Classificações	Descrição
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.

Classificações	Descrição
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.16.1 do Amazon EMR

Versões 5.16.1 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.16.1	emr-5.16.0	emr-5.15.1	emr-5.15.0
AWS SDK for Java	1.11.336	1.11.336	1.11.33	1.11.33
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.5.0	1.5.0	1.4.2	1.4.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.4	1.4.4	1.4.4	1.4.4

	emr-5.16.1	emr-5.16.0	emr-5.15.1	emr-5.15.0
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.3	2.8.3
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.4.0	0.4.0
MXNet	1.2.0	1.2.0	1.1.0	1.1.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.203	0.203	0.194	0.194
Spark	2.3.1	2.3.1	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.16.1	emr-5.16.0	emr-5.15.1	emr-5.15.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

Notas da versão 5.16.1

Esta é uma versão de patch para adicionar a autenticação AWS Signature versão 4 para solicitações ao Amazon S3. Todas as aplicações e componentes são os mesmos que os da versão anterior do Amazon EMR.

> **Important**
>
> Nesta versão de lançamento, o Amazon EMR usa o AWS Signature versão 4 exclusivamente para autenticar solicitações para o Amazon S3. Para obter mais informações, consulte Novidades.

Versões de componente 5.16.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.1.0	SDK do Amazon SageMaker Spark
emr-ddb	4.6.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.10.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.25.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.5.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.4-amzn-0	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.4-amzn-0	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.4-amzn-0	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.4-amzn-0	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.8.4-amzn-0	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.4-amzn-0	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.8.4-amzn-0	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.4-amzn-0	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.8.4-amzn-0	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.4-amzn-0	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.4.4	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.4	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.4	HBase cliente de linha de comando.
hbase-rest-server	1.4.4	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.4	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.3-amzn-1	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.3-amzn-1	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	2.3.3-amzn-1	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.3-amzn-1	O cliente da linha de comando do Hive.
hive-hbase	2.3.3-amzn-1	Cliente do Hive-hbase.
hive-metastore-server	2.3.3-amzn-1	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.3-amzn-1	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.2.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	0.8.1	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.5.0-incubating	Interface REST para interagir com o Apache Spark
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.2.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.

Componente	Versão	Descrição
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.0.0	O cliente da linha de comando do Oozie.
oozie-server	5.0.0	O serviço que aceita solicitações de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computacional de Código Aberto.
phoenix-library	4.14.0- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.0- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.203	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.203	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística

Componente	Versão	Descrição
spark-client	2.3.1	Os clientes da linha de comando do Spark.
spark-history-server	2.3.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.3.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.3.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.3	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.12	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.12	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.16.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.16.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop

Classificações	Descrição
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.

Classificações	Descrição
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.

Classificações	Descrição
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.

Classificações	Descrição
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.

Classificações	Descrição
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.

Classificações	Descrição
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.16.0 do Amazon EMR

Versões 5.16.0 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.16.0	emr-5.15.1	emr-5.15.0	emr-5.14.2
AWS SDK for Java	1.11.336	1.11.33	1.11.33	1.11.297
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.5.0	1.4.2	1.4.2	1.4.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.4	1.4.4	1.4.4	1.4.2

	emr-5.16.0	emr-5.15.1	emr-5.15.0	emr-5.14.2
HCatalog	2.3.3	2.3.3	2.3.3	2.3.2
Hadoop	2.8.4	2.8.3	2.8.3	2.8.3
Hive	2.3.3	2.3.3	2.3.3	2.3.2
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.4.0	0.4.0	0.4.0
MXNet	1.2.0	1.1.0	1.1.0	1.1.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	4.3.0
Phoenix	4.14.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.203	0.194	0.194	0.194
Spark	2.3.1	2.3.0	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.16.0	emr-5.15.1	emr-5.15.0	emr-5.14.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.10

Notas da versão 5.16.0

As notas da versão a seguir incluem informações para a versão 5.16.0 do Amazon EMR. As alterações são referentes à versão 5.15.0.

Data da versão inicial: 19 de julho de 2018

Atualizações

- Hadoop 2.8.4
- Flink 1.5.0
- Livy 0.5.0
- MXNet 1.2.0
- Phoenix 4.14.0
- Presto 0.203
- Spark 2.3.1
- AWS SDK para Java 1.11.336
- CUDA 9.2
- Driver JDBC do Redshift, 1.2.15.1025

Alterações, melhorias e problemas resolvidos

- HBase
  - Backport do HBASE-20723
- Presto
  - Alterações na configuração para oferecer suporte à autenticação LDAP. Para obter mais informações, consulte Usar autenticação LDAP para o Presto no Amazon EMR.

- Spark
  - A versão 2.3.1 do Apache Spark, disponível desde a versão 5.16.0 do Amazon EMR, aborda CVE-2018-8024 e CVE-2018-1334. Recomendamos que você migre as versões anteriores do Spark para a versão 2.3.1 ou posteriores.

Problemas conhecidos

- Essa versão não é compatível com os tipos de instância c1.medium ou m1.small. Os clusters que usam qualquer um desses tipos de instância não são iniciados. Como solução alternativa, especifique um tipo de instância diferente ou use uma versão diferente.
- Quando você cria um cluster kerberizado com o Livy instalado, o Livy apresenta falha com um erro em que a autenticação simples não está habilitada. A reinicialização do servidor do Livy resolve o problema. Como solução alternativa, adicione uma etapa durante a criação do cluster que execute `sudo restart livy-server` no nó primário.
- Depois que o nó primário for reinicializado ou o controlador de instância for reiniciado, as CloudWatch métricas não serão coletadas e o recurso de escalabilidade automática não estará disponível nas versões 5.14.0, 5.15.0 ou 5.16.0 do Amazon EMR. Esse problema foi corrigido na versão 5.17.0 do Amazon EMR.

Versões de componente 5.16.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.1.0	SDK do Amazon SageMaker Spark
emr-ddb	4.6.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.10.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.25.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.5.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.4-amzn-0	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.4-amzn-0	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.4-amzn-0	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.4-amzn-0	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.8.4-amzn-0	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.4-amzn-0	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.8.4-amzn-0	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.4-amzn-0	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.8.4-amzn-0	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.4-amzn-0	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.4.4	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.4	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.4	HBase cliente de linha de comando.
hbase-rest-server	1.4.4	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.4	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.3-amzn-1	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.3-amzn-1	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	2.3.3-amzn-1	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.3-amzn-1	O cliente da linha de comando do Hive.
hive-hbase	2.3.3-amzn-1	Cliente do Hive-hbase.
hive-metastore-server	2.3.3-amzn-1	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.3-amzn-1	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.2.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	0.8.1	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.5.0-incubating	Interface REST para interagir com o Apache Spark
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.2.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.

Componente	Versão	Descrição
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.2.88	Drivers NVIDIA e toolkit CUDA
oozie-client	5.0.0	O cliente da linha de comando do Oozie.
oozie-server	5.0.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.14.0- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.14.0- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.203	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.203	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística

Componente	Versão	Descrição
spark-client	2.3.1	Os clientes da linha de comando do Spark.
spark-history-server	2.3.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.3.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.3.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.3	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.12	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.12	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.16.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.16.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop

Classificações	Descrição
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.

Classificações	Descrição
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.

Classificações	Descrição
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.

Classificações	Descrição
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-password-authenticator	Altere os valores no arquivo password-authenticator.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.

Classificações	Descrição
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.

Classificações	Descrição
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.15.1 do Amazon EMR

Versões 5.15.1 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.15.1	emr-5.15.0	emr-5.14.2	emr-5.14.1
AWS SDK for Java	1.11.33	1.11.33	1.11.297	1.11.297
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.2	1.4.2	1.4.2	1.4.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.4	1.4.4	1.4.2	1.4.2

	emr-5.15.1	emr-5.15.0	emr-5.14.2	emr-5.14.1
HCatalog	2.3.3	2.3.3	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.3	2.3.3	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.1.0	1.1.0	1.1.0	1.1.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.194	0.194
Spark	2.3.0	2.3.0	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.15.1	emr-5.15.0	emr-5.14.2	emr-5.14.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.10	3.4.10

Notas da versão 5.15.1

Esta é uma versão de patch para adicionar a autenticação AWS Signature versão 4 para solicitações ao Amazon S3. Todas as aplicações e componentes são os mesmos que os da versão anterior do Amazon EMR.

> **Important**
>
> Nesta versão de lançamento, o Amazon EMR usa o AWS Signature versão 4 exclusivamente para autenticar solicitações para o Amazon S3. Para obter mais informações, consulte Novidades.

Versões de componente 5.15.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.0.1	SDK do Amazon SageMaker Spark
emr-ddb	4.5.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.10.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.24.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.4.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.3-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.3-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.3-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.3-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.8.3-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.3-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.3-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.8.3-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.3-amzn-1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.4.4	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.4	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.4	HBase cliente de linha de comando.
hbase-rest-server	1.4.4	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.4	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.3-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.3-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	2.3.3-amzn-0	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.3-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.3.3-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.3.3-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.3-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.2.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	0.8.1	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.4.0-incubating	Interface REST para interagir com o Apache Spark
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.1.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.

Componente	Versão	Descrição
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.1.85	Drivers NVIDIA e toolkit CUDA
oozie-client	5.0.0	O cliente da linha de comando do Oozie.
oozie-server	5.0.0	O serviço que aceita solicitações de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computacional de Código Aberto.
phoenix-library	4.13.0- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.13.0- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.194	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.194	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística

Componente	Versão	Descrição
spark-client	2.3.0	Os clientes da linha de comando do Spark.
spark-history-server	2.3.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.3.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.3.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.3	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.12	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.12	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.15.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.15.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop

Classificações	Descrição
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.

Classificações	Descrição
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.

Classificações	Descrição
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.

Classificações	Descrição
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra .properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.

Classificações	Descrição
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.

Classificações	Descrição
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.15.0 do Amazon EMR

Versões 5.15.0 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.15.0	emr-5.14.2	emr-5.14.1	emr-5.14.0
AWS SDK for Java	1.11.33	1.11.297	1.11.297	1.11.297
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.2	1.4.2	1.4.2	1.4.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.4	1.4.2	1.4.2	1.4.2

	emr-5.15.0	emr-5.14.2	emr-5.14.1	emr-5.14.0
HCatalog	2.3.3	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.3	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.2.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.1.0	1.1.0	1.1.0	1.1.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.194	0.194
Spark	2.3.0	2.3.0	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.15.0	emr-5.14.2	emr-5.14.1	emr-5.14.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.10	3.4.10	3.4.10

Notas da versão 5.15.0

As notas da versão a seguir incluem informações para a versão 5.15.0 do Amazon EMR. As alterações são referentes à versão 5.14.0.

Data da versão inicial: 21 de junho de 2018

Atualizações

- Atualizado HBase para 1.4.4
- Atualizado Hive para 2.3.3
- Atualizado Hue para 4.2.0
- Atualizado Oozie para 5.0.0
- Atualizado Zookeeper para 3.4.12
- AWS SDK atualizado para 1.11.333

Alterações, melhorias e problemas resolvidos

- Hive
  - Backport do HIVE-18069
- Hue
  - Atualizado o Hue para se autenticar corretamente com o Livy quando o Kerberos está habilitado. Agora, o Livy é compatível quando usa o Kerberos com o Amazon EMR.
- JupyterHub
  - Atualizado JupyterHub para que o Amazon EMR instale bibliotecas de clientes LDAP por padrão.
  - Corrigido um erro no script que gera certificados autoassinados.

Problemas conhecidos

- Essa versão não é compatível com os tipos de instância c1.medium ou m1.small. Os clusters que usam qualquer um desses tipos de instância não são iniciados. Como solução alternativa, especifique um tipo de instância diferente ou use uma versão diferente.
- Depois que o nó primário for reinicializado ou o controlador de instância for reiniciado, as CloudWatch métricas não serão coletadas e o recurso de escalabilidade automática não estará disponível nas versões 5.14.0, 5.15.0 ou 5.16.0 do Amazon EMR. Esse problema foi corrigido na versão 5.17.0 do Amazon EMR.

Versões de componente 5.15.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.0.1	SDK do Amazon SageMaker Spark
emr-ddb	4.5.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.

Componente	Versão	Descrição
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.10.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.24.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.4.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.3-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.

Componente	Versão	Descrição
hadoop-hdfs-datanode	2.8.3-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.3-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.3-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.8.3-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.3-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.3-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.8.3-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.3-amzn-1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.

Componente	Versão	Descrição
hbase-hmaster	1.4.4	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.4	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.4	HBase cliente de linha de comando.
hbase-rest-server	1.4.4	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.4	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.3-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.3-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.3.3-amzn-0	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.3-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.3.3-amzn-0	Cliente do Hive-hbase.

Componente	Versão	Descrição
hive-metastore-server	2.3.3-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.3-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.2.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	0.8.1	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.4.0-incubating	Interface REST para interagir com o Apache Spark
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.1.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.1.85	Drivers NVIDIA e toolkit CUDA
oozie-client	5.0.0	O cliente da linha de comando do Oozie.

Componente	Versão	Descrição
oozie-server	5.0.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.13.0- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.13.0- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.194	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.194	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística
spark-client	2.3.0	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	2.3.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.3.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.3.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.3	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.12	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.12	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.15.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.15.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop

Classificações	Descrição
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.

Classificações	Descrição
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.

Classificações	Descrição
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.

Classificações	Descrição
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.

Classificações	Descrição
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.

Classificações	Descrição
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.14.2 do Amazon EMR

Versões 5.14.2 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.14.2	emr-5.14.1	emr-5.14.0	emr-5.13.1
AWS SDK for Java	1.11.297	1.11.297	1.11.297	1.11.297
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.2	1.4.2	1.4.2	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.2	1.4.2	1.4.2	1.4.2

	emr-5.14.2	emr-5.14.1	emr-5.14.0	emr-5.13.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.1.0	1.1.0	1.1.0	1.0.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.194	0.194
Spark	2.3.0	2.3.0	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.14.2	emr-5.14.1	emr-5.14.0	emr-5.13.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.14.2

Esta é uma versão de patch para adicionar a autenticação AWS Signature versão 4 para solicitações ao Amazon S3. Todas as aplicações e componentes são os mesmos que os da versão anterior do Amazon EMR.

> **Important**
>
> Nesta versão de lançamento, o Amazon EMR usa o AWS Signature versão 4 exclusivamente para autenticar solicitações para o Amazon S3. Para obter mais informações, consulte Novidades.

Versões de componente 5.14.2

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.0.1	SDK do Amazon SageMaker Spark
emr-ddb	4.5.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.10.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.23.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.4.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.3-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.3-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.3-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.3-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.8.3-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.3-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.3-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.8.3-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.3-amzn-1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.4.2	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.2	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.2	HBase cliente de linha de comando.
hbase-rest-server	1.4.2	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.2	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.2-amzn-2	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.2-amzn-2	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	2.3.2-amzn-2	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.2-amzn-2	O cliente da linha de comando do Hive.
hive-hbase	2.3.2-amzn-2	Cliente do Hive-hbase.
hive-metastore-server	2.3.2-amzn-2	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.2-amzn-2	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.1.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	0.8.1	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.4.0-incubating	Interface REST para interagir com o Apache Spark
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.1.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.

Componente	Versão	Descrição
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.1.85	Drivers NVIDIA e toolkit CUDA
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitações de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computacional de Código Aberto.
phoenix-library	4.13.0- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.13.0- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.194	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.194	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística

Componente	Versão	Descrição
spark-client	2.3.0	Os clientes da linha de comando do Spark.
spark-history-server	2.3.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.3.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.3.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.3	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.14.2

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.14.2

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop

Classificações	Descrição
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.

Classificações	Descrição
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.

Classificações	Descrição
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.

Classificações	Descrição
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.

Classificações	Descrição
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.

Classificações	Descrição
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.14.1 do Amazon EMR

Versões 5.14.1 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.14.1	emr-5.14.0	emr-5.13.1	emr-5.13.0
AWS SDK for Java	1.11.297	1.11.297	1.11.297	1.11.297
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.2	1.4.2	1.4.0	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.2	1.4.2	1.4.2	1.4.2

	emr-5.14.1	emr-5.14.0	emr-5.13.1	emr-5.13.0
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.1.0	1.1.0	1.0.0	1.0.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.194	0.194
Spark	2.3.0	2.3.0	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.14.1	emr-5.14.0	emr-5.13.1	emr-5.13.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.14.1

As notas da versão a seguir incluem informações para a versão 5.14.1 do Amazon EMR. As alterações são referentes à versão 5.14.0.

Data da versão inicial: 17 de outubro de 2018

Atualizada a AMI padrão para Amazon EMR para abordar possíveis vulnerabilidades de segurança.

Versões de componente 5.14.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.0.1	SDK do Amazon SageMaker Spark

Componente	Versão	Descrição
emr-ddb	4.5.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.10.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.23.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.4.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.

Componente	Versão	Descrição
hadoop-client	2.8.3-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.3-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.3-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.3-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.8.3-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.3-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.8.3-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.3-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.8.3-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.

Componente	Versão	Descrição
hadoop-yarn-timeline-server	2.8.3-amzn-1	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.4.2	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.2	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.2	HBase cliente de linha de comando.
hbase-rest-server	1.4.2	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.2	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.2-amzn-2	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.2-amzn-2	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.3.2-amzn-2	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.

Componente	Versão	Descrição
hive-client	2.3.2-amzn-2	O cliente da linha de comando do Hive.
hive-hbase	2.3.2-amzn-2	Cliente do Hive-hbase.
hive-metastore-server	2.3.2-amzn-2	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.2-amzn-2	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.1.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	0.8.1	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.4.0-incubating	Interface REST para interagir com o Apache Spark
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.1.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.1.85	Drivers NVIDIA e toolkit CUDA

Componente	Versão	Descrição
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computaci onal de Código Aberto.
phoenix-library	4.13.0- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.13.0- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.194	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.194	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística
spark-client	2.3.0	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	2.3.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.3.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.3.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.3	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.14.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.14.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop

Classificações	Descrição
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.

Classificações	Descrição
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.

Classificações	Descrição
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.

Classificações	Descrição
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.

Classificações	Descrição
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.

Classificações	Descrição
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.14.0 do Amazon EMR

Versões 5.14.0 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.14.0	emr-5.13.1	emr-5.13.0	emr-5.12.3
AWS SDK for Java	1.11.297	1.11.297	1.11.297	1.11.267
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.2	1.4.0	1.4.0	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.2	1.4.2	1.4.2	1.4.0

	emr-5.14.0	emr-5.13.1	emr-5.13.0	emr-5.12.3
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.1.0	1.0.0	1.0.0	1.0.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.194	0.188
Spark	2.3.0	2.3.0	2.3.0	2.2.1
Sqoop	1.4.7	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.14.0	emr-5.13.1	emr-5.13.0	emr-5.12.3
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.14.0

As notas da versão a seguir incluem informações para a versão 5.14.0 do Amazon EMR. As alterações são referentes à versão 5.13.0.

Data da versão inicial: 4 de junho de 2018

Atualizações

- Atualizado Apache Flink para 1.4.2
- Apache MXnet atualizado para 1.1.0
- Atualizado Apache Sqoop para 1.4.7

Novos recursos

- JupyterHub Suporte adicionado. Para obter mais informações, consulte JupyterHub.

Alterações, melhorias e problemas resolvidos

- EMRFS
  - A string userAgent nas solicitações ao Amazon S3 foi atualizada para conter as informações de usuário e grupo da entidade principal invocadora. Isso pode ser usado com AWS CloudTrail registros para um rastreamento de solicitações mais abrangente.
- HBase
  - Incluído o HBASE-20447, que aborda um problema que poderia causar falhas de cache, especialmente com regiões divididas.
- MXnet

  - Adicionadas bibliotecas OpenCV.
- Spark
  - Quando o Spark grava arquivos Parquet em um local do Amazon S3 usando o EMRFS, FileOutputCommitter o algoritmo foi atualizado para usar a versão 2 em vez da versão 1. Isso reduz o número de renomeações, o que melhora a performance da aplicação. Essa alteração não afeta:
    - Aplicações diferentes do Spark.
    - Aplicativos que gravam em outros sistemas de arquivos, como o HDFS (que ainda usa a versão 1 do FileOutputCommitter).
    - Aplicações que usam outros formatos de saída, como texto ou csv, que já usam a gravação direta do EMRFS.

Problemas conhecidos

- JupyterHub
  - O uso de classificações de configuração para configurar JupyterHub notebooks Jupyter individuais ao criar um cluster não é suportado. Edite manualmente o arquivo jupyterhub_config.py e os arquivos jupyter_notebook_config.py para cada usuário. Para obter mais informações, consulte Configurando JupyterHub.
  - JupyterHub falha ao iniciar em clusters dentro de uma sub-rede privada, falhando com a mensagem. `Error: ENOENT: no such file or directory, open '/etc/jupyter/conf/server.crt'` Isso é causado por um erro no script que gera certificados autoassinados. Use a solução alternativa a seguir para gerar certificados autoassinados. Todos os comandos são executados enquanto estão conectados ao nó primário.
    1. Copie o script de geração de certificados do contêiner para o nó primário:

       ```
       sudo docker cp jupyterhub:/tmp/gen_self_signed_cert.sh ./
       ```

    2. Use um editor de texto para alterar a linha 23 e mudar o nome de host público para o nome deo host local, conforme mostrado abaixo:

       ```
       local hostname=$(curl -s $EC2_METADATA_SERVICE_URI/local-hostname)
       ```

    3. Execute o script para gerar certificados autoassinados:

```
sudo bash ./gen_self_signed_cert.sh
```

4. Mova os arquivos de certificado que o script gera para o diretório `/etc/jupyter/conf/`:

```
sudo mv /tmp/server.crt /tmp/server.key /etc/jupyter/conf/
```

Você pode acessar `tail` o `jupyter.log` arquivo para verificar se ele JupyterHub foi reiniciado e está retornando um código de resposta 200. Por exemplo:

```
tail -f /var/log/jupyter/jupyter.log
```

Essa ação deve retornar uma resposta semelhante à seguinte:

```
# [I 2018-06-14 18:56:51.356 JupyterHub app:1581] JupyterHub is now running at
 https://:9443/
# 19:01:51.359 - info: [ConfigProxy] 200 GET /api/routes
```

- Depois que o nó primário for reinicializado ou o controlador de instância for reiniciado, as CloudWatch métricas não serão coletadas e o recurso de escalabilidade automática não estará disponível nas versões 5.14.0, 5.15.0 ou 5.16.0 do Amazon EMR. Esse problema foi corrigido na versão 5.17.0 do Amazon EMR.

Versões de componente 5.14.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.0.1	SDK do Amazon SageMaker Spark
emr-ddb	4.5.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.10.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.23.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.4.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.3-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.3-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.3-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.3-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.8.3-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.3-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.3-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.8.3-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.3-amzn-1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.4.2	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.2	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.2	HBase cliente de linha de comando.
hbase-rest-server	1.4.2	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.2	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.2-amzn-2	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.2-amzn-2	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	2.3.2-amzn-2	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.2-amzn-2	O cliente da linha de comando do Hive.
hive-hbase	2.3.2-amzn-2	Cliente do Hive-hbase.
hive-metastore-server	2.3.2-amzn-2	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.2-amzn-2	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.1.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
jupyterhub	0.8.1	Servidor de vários usuários para blocos de anotações Jupyter
livy-server	0.4.0-incubating	Interface REST para interagir com o Apache Spark
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.1.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.

Componente	Versão	Descrição
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.1.85	Drivers NVIDIA e toolkit CUDA
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitações de fluxo de trabalho do Oozie.
opencv	3.4.0	Biblioteca de Visão Computacional de Código Aberto.
phoenix-library	4.13.0- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.13.0- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.194	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.194	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística

Componente	Versão	Descrição
spark-client	2.3.0	Os clientes da linha de comando do Spark.
spark-history-server	2.3.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.3.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.3.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.7	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.3	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.14.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.14.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
container-log4j	Altere os valores no arquivo container-log4j.properties do YARN do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop

Classificações	Descrição
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.

Classificações	Descrição
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.

Classificações	Descrição
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
jupyter-notebook-conf	Alterar os valores no arquivo jupyter_notebook_config.py do Notebook Jupyter.
jupyter-hub-conf	Altere os valores no JupyterHubs arquivo jupyterhub_config.py do.
jupyter-sparkmagic-conf	Altere os valores no arquivo config.json do Sparkmagic.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.

Classificações	Descrição
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra .properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.

Classificações	Descrição
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.

Classificações	Descrição
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.13.1 do Amazon EMR

Versões 5.13.1 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, e ZooKeeper.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.13.1	emr-5.13.0	emr-5.12.3	emr-5.12.2
AWS SDK for Java	1.11.297	1.11.297	1.11.267	1.11.267
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.0	1.4.0	1.4.0	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.2	1.4.2	1.4.0	1.4.0

	emr-5.13.1	emr-5.13.0	emr-5.12.3	emr-5.12.2
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.0.0	1.0.0	1.0.0	1.0.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.188	0.188
Spark	2.3.0	2.3.0	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.13.1	emr-5.13.0	emr-5.12.3	emr-5.12.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.13.1

Esta é uma versão de patch para adicionar a autenticação AWS Signature versão 4 para solicitações ao Amazon S3. Todas as aplicações e componentes são os mesmos que os da versão anterior do Amazon EMR.

> **Important**
>
> Nesta versão de lançamento, o Amazon EMR usa o AWS Signature versão 4 exclusivamente para autenticar solicitações para o Amazon S3. Para obter mais informações, consulte Novidades.

Versões de componente 5.13.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.0.1	SDK do Amazon SageMaker Spark
emr-ddb	4.5.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.10.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.22.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.4.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.3-amzn-0	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.3-amzn-0	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.3-amzn-0	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.3-amzn-0	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.8.3-amzn-0	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.3-amzn-0	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-0	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.3-amzn-0	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.8.3-amzn-0	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.3-amzn-0	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.4.2	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.2	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.2	HBase cliente de linha de comando.
hbase-rest-server	1.4.2	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.2	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.2-amzn-2	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.2-amzn-2	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	2.3.2-amzn-2	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.2-amzn-2	O cliente da linha de comando do Hive.
hive-hbase	2.3.2-amzn-2	Cliente do Hive-hbase.
hive-metastore-server	2.3.2-amzn-2	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.2-amzn-2	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.1.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
livy-server	0.4.0-incubating	Interface REST para interagir com o Apache Spark
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.0.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.1.85	Drivers NVIDIA e toolkit CUDA

Componente	Versão	Descrição
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.13.0- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.13.0- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.194	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.194	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística
spark-client	2.3.0	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	2.3.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.3.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.3.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.3	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.13.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.13.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop

Classificações	Descrição
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.

Classificações	Descrição
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.

Classificações	Descrição
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.

Classificações	Descrição
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.

Classificações	Descrição
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.

Classificações	Descrição
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.13.0 do Amazon EMR

Versões 5.13.0 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, e ZooKeeper.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR

- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.13.0	emr-5.12.3	emr-5.12.2	emr-5.12.1
AWS SDK for Java	1.11.297	1.11.267	1.11.267	1.11.267
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.0	1.4.0	1.4.0	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.2	1.4.0	1.4.0	1.4.0
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-

	emr-5.13.0	emr-5.12.3	emr-5.12.2	emr-5.12.1
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.0.0	1.0.0	1.0.0	1.0.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.188	0.188	0.188
Spark	2.3.0	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.13.0

As notas da versão a seguir incluem informações para a versão 5.13.0 do Amazon EMR. As alterações são referentes à versão 5.12.0.

Atualizações

- Atualizado Spark para 2.3.0
- Atualizado HBase para 1.4.2

- Atualizado Presto para 0.194
- Atualizado para AWS SDK para Java 1.11.297

Alterações, melhorias e problemas resolvidos

- Hive
  - Backport do HIVE-15436. Hive aprimorado APIs para retornar somente visualizações.

Problemas conhecidos

- MXNet atualmente não tem bibliotecas OpenCV.

Versões de componente 5.13.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.0.1	SDK do Amazon SageMaker Spark
emr-ddb	4.5.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.

Componente	Versão	Descrição
emr-goodies	2.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.10.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.22.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.4.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.3-amzn-0	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.

Componente	Versão	Descrição
hadoop-hdfs-datanode	2.8.3-amzn-0	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.3-amzn-0	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.3-amzn-0	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.8.3-amzn-0	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.3-amzn-0	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-0	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.3-amzn-0	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.8.3-amzn-0	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.3-amzn-0	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.

Componente	Versão	Descrição
hbase-hmaster	1.4.2	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.2	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.2	HBase cliente de linha de comando.
hbase-rest-server	1.4.2	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.2	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.2-amzn-2	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.2-amzn-2	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.3.2-amzn-2	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.2-amzn-2	O cliente da linha de comando do Hive.
hive-hbase	2.3.2-amzn-2	Cliente do Hive-hbase.

Componente	Versão	Descrição
hive-metastore-server	2.3.2-amzn-2	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.2-amzn-2	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.1.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
livy-server	0.4.0-incubating	Interface REST para interagir com o Apache Spark
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.0.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.1.85	Drivers NVIDIA e toolkit CUDA
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.

Componente	Versão	Descrição
phoenix-library	4.13.0- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.13.0- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.194	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.194	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
r	3.4.1	O projeto R para computação estatística
spark-client	2.3.0	Os clientes da linha de comando do Spark.
spark-history-server	2.3.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.3.0	O mecanismo de execução em memória para o YARN.

Componente	Versão	Descrição
spark-yarn-slave	2.3.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.3	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.13.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.13.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.

Classificações	Descrição
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive- site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.pro perties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.pr operties do Hive.

Classificações	Descrição
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.proper ties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.x ml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properti es do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.

Classificações	Descrição
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.

Classificações	Descrição
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.

Classificações	Descrição
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.

Classificações	Descrição
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.12.3 do Amazon EMR

Versões 5.12.3 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, e ZooKeeper.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.12.3	emr-5.12.2	emr-5.12.1	emr-5.12.0
AWS SDK for Java	1.11.267	1.11.267	1.11.267	1.11.267
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.0	1.4.0	1.4.0	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.0	1.4.0	1.4.0	1.4.0
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.0.0	1.0.0	1.0.0	1.0.0

	emr-5.12.3	emr-5.12.2	emr-5.12.1	emr-5.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.188	0.188	0.188	0.188
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.12.3

Esta é uma versão de patch para adicionar a autenticação AWS Signature versão 4 para solicitações ao Amazon S3. Todas as aplicações e componentes são os mesmos que os da versão anterior do Amazon EMR.

> **Important**
>
> Nesta versão de lançamento, o Amazon EMR usa o AWS Signature versão 4 exclusivamente para autenticar solicitações para o Amazon S3. Para obter mais informações, consulte Novidades.

Versões de componente 5.12.3

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.0.1	SDK do Amazon SageMaker Spark
emr-ddb	4.5.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.9.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.21.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.

Componente	Versão	Descrição
flink-client	1.4.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.3-amzn-0	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.3-amzn-0	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.3-amzn-0	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.3-amzn-0	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.8.3-amzn-0	O endpoint de HTTP para as operações do HDFS.

Componente	Versão	Descrição
hadoop-kms-server	2.8.3-amzn-0	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-0	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.3-amzn-0	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.8.3-amzn-0	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.3-amzn-0	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.4.0	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.0	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.0	HBase cliente de linha de comando.

Componente	Versão	Descrição
hbase-rest-server	1.4.0	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.0	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.2-amzn-1	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.2-amzn-1	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.3.2-amzn-1	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.2-amzn-1	O cliente da linha de comando do Hive.
hive-hbase	2.3.2-amzn-1	Cliente do Hive-hbase.
hive-metastore-server	2.3.2-amzn-1	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.2-amzn-1	O serviço que aceita as consultas do Hive como solicitações da web.

Componente	Versão	Descrição
hue-server	4.1.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
livy-server	0.4.0-incubating	Interface REST para interagir com o Apache Spark
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.0.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.1.85	Drivers NVIDIA e toolkit CUDA
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.13.0- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.13.0- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica

Componente	Versão	Descrição
presto-coordinator	0.188	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.188	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
spark-client	2.2.1	Os clientes da linha de comando do Spark.
spark-history-server	2.2.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.2.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.2.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.

Componente	Versão	Descrição
zeppelin-server	0.7.3	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.12.3

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.12.3

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.

Classificações	Descrição
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.

Classificações	Descrição
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.

Classificações	Descrição
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.

Classificações	Descrição
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.

Classificações	Descrição
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.

Classificações	Descrição
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.

Classificações	Descrição
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.12.2 do Amazon EMR

Versões 5.12.2 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, e ZooKeeper.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.12.2	emr-5.12.1	emr-5.12.0	emr-5.11.4
AWS SDK for Java	1.11.267	1.11.267	1.11.267	1.11.238
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-5.12.2	emr-5.12.1	emr-5.12.0	emr-5.11.4
Flink	1.4.0	1.4.0	1.4.0	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.0	1.4.0	1.4.0	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.0.0	1.0.0	1.0.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.188	0.188	0.188	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6

	emr-5.12.2	emr-5.12.1	emr-5.12.0	emr-5.11.4
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.12.2

As notas da versão a seguir incluem informações para a versão 5.12.2 do Amazon EMR. As alterações são referentes à versão 5.12.1.

Data da versão inicial: 29 de agosto de 2018

Alterações, melhorias e problemas resolvidos

- Esta versão aborda uma possível vulnerabilidade de segurança.

Versões de componente 5.12.2

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.0.1	SDK do Amazon SageMaker Spark
emr-ddb	4.5.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.9.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.21.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.4.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.3-amzn-0	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.3-amzn-0	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.3-amzn-0	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.3-amzn-0	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.8.3-amzn-0	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.3-amzn-0	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-0	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.3-amzn-0	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.8.3-amzn-0	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.3-amzn-0	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.4.0	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.0	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.0	HBase cliente de linha de comando.
hbase-rest-server	1.4.0	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.0	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.2-amzn-1	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.2-amzn-1	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	2.3.2-amzn-1	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.2-amzn-1	O cliente da linha de comando do Hive.
hive-hbase	2.3.2-amzn-1	Cliente do Hive-hbase.
hive-metastore-server	2.3.2-amzn-1	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.2-amzn-1	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.1.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
livy-server	0.4.0-incubating	Interface REST para interagir com o Apache Spark
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.0.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.1.85	Drivers NVIDIA e toolkit CUDA

Componente	Versão	Descrição
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.13.0- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.13.0- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.188	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.188	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
spark-client	2.2.1	Os clientes da linha de comando do Spark.
spark-history-server	2.2.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.

Componente	Versão	Descrição
spark-on-yarn	2.2.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.2.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.3	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.12.2

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.12.2

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.

Classificações	Descrição
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive- site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.pro perties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.pr operties do Hive.

Classificações	Descrição
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.proper ties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.x ml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properti es do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.

Classificações	Descrição
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.

Classificações	Descrição
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.

Classificações	Descrição
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.

Classificações	Descrição
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.12.1 do Amazon EMR

Versões 5.12.1 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, e ZooKeeper.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.12.1	emr-5.12.0	emr-5.11.4	emr-5.11.3
AWS SDK for Java	1.11.267	1.11.267	1.11.238	1.11.238
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.0	1.4.0	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.0	1.4.0	1.3.1	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.0.0	1.0.0	0.12.0	0.12.0

	emr-5.12.1	emr-5.12.0	emr-5.11.4	emr-5.11.3
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.188	0.188	0.187	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.12.1

As notas da versão a seguir incluem informações para a versão 5.12.1 do Amazon EMR. As alterações são referentes à versão 5.12.0.

Data da versão inicial: 29 de março de 2018

Alterações, melhorias e problemas resolvidos

- Atualizado o kernel do Amazon Linux da AMI padrão do Amazon Linux para Amazon EMR para abordar possíveis vulnerabilidades.

Versões de componente 5.12.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.0.1	SDK do Amazon SageMaker Spark
emr-ddb	4.5.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.9.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.21.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.

Componente	Versão	Descrição
flink-client	1.4.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.3-amzn-0	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.3-amzn-0	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.3-amzn-0	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.8.3-amzn-0	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.8.3-amzn-0	O endpoint de HTTP para as operações do HDFS.

Componente	Versão	Descrição
hadoop-kms-server	2.8.3-amzn-0	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-0	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.3-amzn-0	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.8.3-amzn-0	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.3-amzn-0	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.4.0	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.4.0	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.0	HBase cliente de linha de comando.

Componente	Versão	Descrição
hbase-rest-server	1.4.0	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.0	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.2-amzn-1	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.2-amzn-1	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.3.2-amzn-1	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.2-amzn-1	O cliente da linha de comando do Hive.
hive-hbase	2.3.2-amzn-1	Cliente do Hive-hbase.
hive-metastore-server	2.3.2-amzn-1	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.2-amzn-1	O serviço que aceita as consultas do Hive como solicitações da web.

Componente	Versão	Descrição
hue-server	4.1.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
livy-server	0.4.0-incubating	Interface REST para interagir com o Apache Spark
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.0.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.1.85	Drivers NVIDIA e toolkit CUDA
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.13.0- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.13.0- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica

Componente	Versão	Descrição
presto-coordinator	0.188	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.188	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
spark-client	2.2.1	Os clientes da linha de comando do Spark.
spark-history-server	2.2.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.2.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.2.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.

Componente	Versão	Descrição
zeppelin-server	0.7.3	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.12.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.12.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.

Classificações	Descrição
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.

Classificações	Descrição
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.

Classificações	Descrição
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.

Classificações	Descrição
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.

Classificações	Descrição
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.

Classificações	Descrição
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.

Classificações	Descrição
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.12.0 do Amazon EMR

Versões da aplicação 5.12.0

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, e ZooKeeper.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.12.0	emr-5.11.4	emr-5.11.3	emr-5.11.2
AWS SDK for Java	1.11.267	1.11.238	1.11.238	1.11.238
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-5.12.0	emr-5.11.4	emr-5.11.3	emr-5.11.2
Flink	1.4.0	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.0	1.3.1	1.3.1	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.0.0	0.12.0	0.12.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.188	0.187	0.187	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6

	emr-5.12.0	emr-5.11.4	emr-5.11.3	emr-5.11.2
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.12.0

As notas da versão a seguir incluem informações para a versão 5.12.0 do Amazon EMR. As alterações são referentes à versão 5.11.1.

Atualizações

- AWS SDK para Java 1.11.238 ⇒ 1.11.267. Para obter mais informações, consulte o AWS SDK for Java Change Log on GitHub.
- Hadoop 2.7.3 ⇒ 2.8.3. Para obter mais informações, consulte Versões do Apache Hadoop.
- Flink 1.3.2 ⇒ 1.4.0. Para obter mais informações, consulte o Anúncio de versão do Apache Flink 1.4.0.
- HBase 1.3.1 ⇒ 1.4.0. Para obter mais informações, consulte o anúncio HBase de lançamento.
- Hue 4.0.1 ⇒ 4.1.0. Para obter mais informações, consulte as Notas de versão.
- MxNet 0,12,0 ⇒ 1,0,0. Para obter mais informações, consulte o MXNet Change Log on GitHub.
- Presto 0.187 ⇒ 0.188. Para obter mais informações, consulte as Notas de versão.

Alterações, melhorias e problemas resolvidos

- Hadoop
  - A propriedade `yarn.resourcemanager.decommissioning.timeout` foi alterada para `yarn.resourcemanager.nodemanager-graceful-decommission-timeout-secs`. Você pode usar essa propriedade para personalizar a redução da escala do cluster

verticalmente. Para obter mais informações, consulte Redução da escala do cluster verticalmente no Guia de gerenciamento do Amazon EMR.

  - A CLI do Hadoop adicionou a opção `-d` ao comando `cp` (copy), que especifica cópia direta. Você pode usar isso para evitar a criação de um arquivo `.COPYING` intermediário, o que torna mais rápida a cópia de dados entre o Amazon S3. Para obter mais informações, consulte HADOOP-12384.

- Pig
  - Adicionada a classificação de configuração `pig-env`, que simplifica a configuração das propriedades do ambiente do Pig. Para obter mais informações, consulte Configurar aplicações.
- Presto
  - Adicionada a classificação de configuração `presto-connector-redshift`, que pode ser usada para configurar os valores no arquivo de configuração `redshift.properties` do Presto. Para obter mais informações, consulte Conector do Redshift na documentação do Presto e Configurar aplicações.
  - O suporte do Presto para EMRFS foi adicionado e é a configuração padrão. As versões anteriores do Amazon EMR usavam o PrestOS3FileSystem, que era a única opção. Para obter mais informações, consulte Configuração do EMRFS e do PrestoS3 FileSystem.

    > **Note**
    >
    > Se você consultar dados subjacentes no Amazon S3 com a versão 5.12.0 do Amazon EMR, poderão ocorrer erros no Presto. Isso acontece porque o Presto não consegue obter valores de classificação de configuração em `emrfs-site.xml`. Como solução alternativa, crie um subdiretório `emrfs` em `usr/lib/presto/plugin/hive-hadoop2/` e crie um link simbólico em `usr/lib/presto/plugin/hive-hadoop2/emrfs` para o arquivo `/usr/share/aws/emr/emrfs/conf/emrfs-site.xml` existente. Em seguida, reinicie o processo presto-server (`sudo presto-server stop` seguido por `sudo presto-server start`).

- Spark
  - SPARK-22036 retroportado: a BigDecimal multiplicação às vezes retorna nula.

Problemas conhecidos

- MXNet não inclui bibliotecas OpenCV.

- O SparkR não está disponível para clusters criados usando uma AMI personalizada porque R não é instalado por padrão nos nós do cluster.

Versões de componente 5.12.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1.0.1	SDK do Amazon SageMaker Spark
emr-ddb	4.5.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.9.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.

Componente	Versão	Descrição
emrfs	2.21.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.4.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.8.3-amzn-0	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.8.3-amzn-0	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.8.3-amzn-0	O cliente de linha de comando e biblioteca do HDFS

Componente	Versão	Descrição
hadoop-hdfs-namenode	2.8.3-amzn-0	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.8.3-amzn-0	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.8.3-amzn-0	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-0	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.8.3-amzn-0	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.8.3-amzn-0	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.8.3-amzn-0	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.4.0	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.

Componente	Versão	Descrição
hbase-region-server	1.4.0	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.4.0	HBase cliente de linha de comando.
hbase-rest-server	1.4.0	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.4.0	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.2-amzn-1	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.2-amzn-1	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.3.2-amzn-1	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.2-amzn-1	O cliente da linha de comando do Hive.
hive-hbase	2.3.2-amzn-1	Cliente do Hive-hbase.
hive-metastore-server	2.3.2-amzn-1	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.

Componente	Versão	Descrição
hive-server2	2.3.2-amzn-1	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.1.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
livy-server	0.4.0-incubating	Interface REST para interagir com o Apache Spark
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	1.0.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.1.85	Drivers NVIDIA e toolkit CUDA
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.13.0- -1,4 HBase	As bibliotecas do phoenix para servidor e cliente

Componente	Versão	Descrição
phoenix-query-server	4.13.0- -1,4 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.188	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.188	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
spark-client	2.2.1	Os clientes da linha de comando do Spark.
spark-history-server	2.2.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.2.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.2.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.

Componente	Versão	Descrição
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.3	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.12.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.12.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.

Classificações	Descrição
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.

Classificações	Descrição
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.

Classificações	Descrição
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.

Classificações	Descrição
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-env	Alterar os valores no ambiente do Pig.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.

Classificações	Descrição
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.

Classificações	Descrição
presto-connector-redshift	Alterar os valores no arquivo de propriedades redshift do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.

Classificações	Descrição
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.11.4 do Amazon EMR

Versões 5.11.4 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, e ZooKeeper.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.11.4	emr-5.11.3	emr-5.11.2	emr-5.11.1
AWS SDK for Java	1.11.238	1.11.238	1.11.238	1.11.238
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8

	emr-5.11.4	emr-5.11.3	emr-5.11.2	emr-5.11.1
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	0.12.0	0.12.0	0.12.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0

	emr-5.11.4	emr-5.11.3	emr-5.11.2	emr-5.11.1
Presto	0.187	0.187	0.187	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.11.4

Esta é uma versão de patch para adicionar a autenticação AWS Signature versão 4 para solicitações ao Amazon S3. Todas as aplicações e componentes são os mesmos que os da versão anterior do Amazon EMR.

> **Important**
>
> Nesta versão de lançamento, o Amazon EMR usa o AWS Signature versão 4 exclusivamente para autenticar solicitações para o Amazon S3. Para obter mais informações, consulte Novidades.

Versões de componente 5.11.4

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a

versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *CommunityVersion*-amzn-*EmrVersion*. O *EmrVersion* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1,0	SDK do Amazon SageMaker Spark
emr-ddb	4.5.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.8.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.20.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.3.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop,

Componente	Versão	Descrição
		juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-6	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-6	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-6	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-6	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-6	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-6	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider

Componente	Versão	Descrição
hadoop-mapred	2.7.3-amzn-6	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-6	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-6	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hadoop-yarn-timeline-server	2.7.3-amzn-6	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.3.1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.3.1	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.3.1	HBase cliente de linha de comando.
hbase-rest-server	1.3.1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.3.1	Serviço que fornece um endpoint Thrift para. HBase

Componente	Versão	Descrição
hcatalog-client	2.3.2-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.2-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.3.2-amzn-0	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.2-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.3.2-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.3.2-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.2-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.0.1	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
livy-server	0.4.0-incubating	Interface REST para interagir com o Apache Spark

Componente	Versão	Descrição
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	0.12.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.0.176	Drivers NVIDIA e toolkit CUDA
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.11.0- -1.3 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.11.0- -1.3 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.187	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.187	O serviço que executa partes de uma consulta.

Componente	Versão	Descrição
pig-client	0.17.0	O cliente da linha de comando do Pig.
spark-client	2.2.1	Os clientes da linha de comando do Spark.
spark-history-server	2.2.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.2.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.2.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.3	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.

Componente	Versão	Descrição
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.11.4

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.11.4

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.

Classificações	Descrição
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.

Classificações	Descrição
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.

Classificações	Descrição
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.

Classificações	Descrição
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.

Classificações	Descrição
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.

Classificações	Descrição
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.11.3 do Amazon EMR

Versões 5.11.3 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, e ZooKeeper.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR

- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.11.3	emr-5.11.2	emr-5.11.1	emr-5.11.0
AWS SDK for Java	1.11.238	1.11.238	1.11.238	1.11.238
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-

	emr-5.11.3	emr-5.11.2	emr-5.11.1	emr-5.11.0
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	0.12.0	0.12.0	0.12.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.187	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.11.3

As notas da versão a seguir incluem informações para a versão 5.11.3 do Amazon EMR. As alterações são referentes à versão 5.11.2.

Data da versão inicial: 18 de julho de 2019

Alterações, melhorias e problemas resolvidos

- Atualizada a AMI padrão do Amazon Linux para Amazon EMR para incluir atualizações de segurança importantes do kernel Linux, incluindo o problema de negação de serviço do TCP SACK (AWS-2019-005).

Versões de componente 5.11.3

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1,0	SDK do Amazon SageMaker Spark
emr-ddb	4.5.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.

Componente	Versão	Descrição
emr-s3-dist-cp	2.8.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.20.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.3.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-6	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-6	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-6	O cliente de linha de comando e biblioteca do HDFS

Componente	Versão	Descrição
hadoop-hdfs-namenode	2.7.3-amzn-6	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-6	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-6	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-6	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-6	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.3-amzn-6	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-6	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.3.1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.

Componente	Versão	Descrição
hbase-region-server	1.3.1	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.3.1	HBase cliente de linha de comando.
hbase-rest-server	1.3.1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.3.1	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.2-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.2-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.3.2-amzn-0	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.2-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.3.2-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.3.2-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.

Componente	Versão	Descrição
hive-server2	2.3.2-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.0.1	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
livy-server	0.4.0-incubating	Interface REST para interagir com o Apache Spark
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	0.12.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.0.176	Drivers NVIDIA e toolkit CUDA
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.11.0- -1.3 HBase	As bibliotecas do phoenix para servidor e cliente

Componente	Versão	Descrição
phoenix-query-server	4.11.0- -1.3 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.187	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.187	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
spark-client	2.2.1	Os clientes da linha de comando do Spark.
spark-history-server	2.2.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.2.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.2.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.

Componente	Versão	Descrição
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.3	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.11.3

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.11.3

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.

Classificações	Descrição
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.

Classificações	Descrição
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.

Classificações	Descrição
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.

Classificações	Descrição
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.

Classificações	Descrição
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.

Classificações	Descrição
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.

Classificações	Descrição
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.11.2 do Amazon EMR

Versões 5.11.2 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, e ZooKeeper.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.11.2	emr-5.11.1	emr-5.11.0	emr-5.10.1
AWS SDK for Java	1.11.238	1.11.238	1.11.238	1.11.21
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-5.11.2	emr-5.11.1	emr-5.11.0	emr-5.10.1
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.1
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	0.12.0	0.12.0	0.12.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.187	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6

	emr-5.11.2	emr-5.11.1	emr-5.11.0	emr-5.10.1
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.11.2

As notas da versão a seguir incluem informações para a versão 5.11.2 do Amazon EMR. As alterações são referentes à versão 5.11.1.

Data da versão inicial: 29 de agosto de 2018

Alterações, melhorias e problemas resolvidos

- Esta versão aborda uma possível vulnerabilidade de segurança.

Versões de componente 5.11.2

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1,0	SDK do Amazon SageMaker Spark
emr-ddb	4.5.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.8.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.20.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.3.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-6	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-6	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-6	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-6	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-6	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-6	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-6	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-6	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.7.3-amzn-6	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-6	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.3.1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.3.1	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.3.1	HBase cliente de linha de comando.
hbase-rest-server	1.3.1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.3.1	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.2-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.2-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	2.3.2-amzn-0	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.2-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.3.2-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.3.2-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.2-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.0.1	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
livy-server	0.4.0-incubating	Interface REST para interagir com o Apache Spark
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	0.12.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.0.176	Drivers NVIDIA e toolkit CUDA

Componente	Versão	Descrição
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.11.0- -1.3 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.11.0- -1.3 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.187	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.187	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
spark-client	2.2.1	Os clientes da linha de comando do Spark.
spark-history-server	2.2.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.

Componente	Versão	Descrição
spark-on-yarn	2.2.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.2.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.3	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.11.2

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.11.2

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.

Classificações	Descrição
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive- site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.pro perties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.pr operties do Hive.

Classificações	Descrição
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.proper ties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.x ml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properti es do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.

Classificações	Descrição
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.

Classificações	Descrição
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.

Classificações	Descrição
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.

Classificações	Descrição
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.11.1 do Amazon EMR

Versões 5.11.1 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, e ZooKeeper.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.11.1	emr-5.11.0	emr-5.10.1	emr-5.10.0
AWS SDK for Java	1.11.238	1.11.238	1.11.21	1.11.21

	emr-5.11.1	emr-5.11.0	emr-5.10.1	emr-5.10.0
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.2	2.3.2	2.3.1	2.3.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.1	2.3.1
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	0.12.0	0.12.0	0.12.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0

	emr-5.11.1	emr-5.11.0	emr-5.10.1	emr-5.10.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.187	0.187
Spark	2.2.1	2.2.1	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.11.1

As notas da versão a seguir incluem informações para a versão 5.11.1 do Amazon EMR. As alterações são referentes à versão 5.8.0 do Amazon EMR.

Data da versão inicial: 22 de janeiro de 2018

Alterações, melhorias e problemas resolvidos

- Atualizado o kernel do Amazon Linux da AMI padrão do Amazon Linux para Amazon EMR para abordar vulnerabilidades associadas à execução especulativa (CVE-2017-5715, CVE-2017-5753 e CVE-2017-5754). Para obter mais informações, consulte https://aws.amazon.com/security/security-bulletins/AWS-2018-013/.

Versões de componente 5.11.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1,0	SDK do Amazon SageMaker Spark
emr-ddb	4.5.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.8.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.20.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.

Componente	Versão	Descrição
flink-client	1.3.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-6	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-6	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-6	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-6	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-6	O endpoint de HTTP para as operações do HDFS.

Componente	Versão	Descrição
hadoop-kms-server	2.7.3-amzn-6	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-6	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-6	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.3-amzn-6	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-6	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.3.1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.3.1	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.3.1	HBase cliente de linha de comando.

Componente	Versão	Descrição
hbase-rest-server	1.3.1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.3.1	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.2-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.2-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.3.2-amzn-0	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.2-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.3.2-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.3.2-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.2-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.

Componente	Versão	Descrição
hue-server	4.0.1	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
livy-server	0.4.0-incubating	Interface REST para interagir com o Apache Spark
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	0.12.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.0.176	Drivers NVIDIA e toolkit CUDA
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.11.0- -1.3 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.11.0- -1.3 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica

Componente	Versão	Descrição
presto-coordinator	0.187	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.187	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
spark-client	2.2.1	Os clientes da linha de comando do Spark.
spark-history-server	2.2.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.2.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.2.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.

Componente	Versão	Descrição
zeppelin-server	0.7.3	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.11.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.11.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.

Classificações	Descrição
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.

Classificações	Descrição
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.

Classificações	Descrição
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.

Classificações	Descrição
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.

Classificações	Descrição
presto-connector-cassandra	Alterar os valores no arquivo cassandra .properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.p roperties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.propert ies do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.con f do Spark.

Classificações	Descrição
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.11.0 do Amazon EMR

Versões 5.11.0 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, e ZooKeeper.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.11.0	emr-5.10.1	emr-5.10.0	emr-5.9.1
AWS SDK for Java	1.11.238	1.11.21	1.11.21	1.11.183
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1

	emr-5.11.0	emr-5.10.1	emr-5.10.0	emr-5.9.1
HCatalog	2.3.2	2.3.1	2.3.1	2.3.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.1	2.3.1	2.3.0
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	0.12.0	0.12.0	0.12.0	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.187	0.184
Spark	2.2.1	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.11.0	emr-5.10.1	emr-5.10.0	emr-5.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.11.0

As notas da versão a seguir incluem informações para a versão 5.11.0 do Amazon EMR. As alterações são referentes à versão 5.10.0.

Atualizações

- Hive 2.3.2
- Spark 2.2.1
- SDK para Java 1.11.238

Novos recursos

- Spark
  - Adicionada a configuração `spark.decommissioning.timeout.threshold`, que aprimora o comportamento de desativação do Spark ao serem usadas instâncias spot. Para obter mais informações, consulte Configurar o comportamento de desativação de nós.
  - Foi adicionado o `aws-sagemaker-spark-sdk` componente ao Spark, que instala o Amazon SageMaker Spark e as dependências associadas para a integração do Spark com a Amazon. SageMaker Você pode usar o Amazon SageMaker Spark para criar pipelines de aprendizado de máquina (ML) do Spark usando os estágios da Amazon. SageMaker Para obter mais informações, consulte o SageMaker Spark Readme sobre GitHub e como usar o Apache Spark com a Amazon SageMaker no Amazon Developer Guide. SageMaker

Problemas conhecidos

- MXNet não inclui bibliotecas OpenCV.

- Por padrão, o Hive 2.3.2 define `hive.compute.query.using.stats=true`. Isso faz com que as consultas obtenham dados de estatísticas existentes em vez de diretamente dos dados, o que pode gerar confusão. Por exemplo, se você tiver uma tabela com `hive.compute.query.using.stats=true` e fizer upload de novos arquivos para a tabela `LOCATION`, a execução de uma consulta `SELECT COUNT(*)` na tabela retornará a contagem das estatísticas, e não selecionará as linhas adicionadas.

  Como alternativa, use o comando `ANALYZE TABLE` para reunir novas estatísticas ou defina `hive.compute.query.using.stats=false`. Para obter mais informações, consulte Estatísticas no Hive na documentação do Apache Hive.

Versões de componente 5.11.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
aws-sagemaker-spark-sdk	1,0	SDK do Amazon SageMaker Spark
emr-ddb	4.5.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.

Componente	Versão	Descrição
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.8.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.20.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.3.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-6	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.

Componente	Versão	Descrição
hadoop-hdfs-datanode	2.7.3-amzn-6	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-6	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-6	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-6	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-6	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-6	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-6	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.3-amzn-6	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-6	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.

Componente	Versão	Descrição
hbase-hmaster	1.3.1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.3.1	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.3.1	HBase cliente de linha de comando.
hbase-rest-server	1.3.1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.3.1	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.2-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.2-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.3.2-amzn-0	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.2-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.3.2-amzn-0	Cliente do Hive-hbase.

Componente	Versão	Descrição
hive-metastore-server	2.3.2-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.2-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.0.1	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
livy-server	0.4.0-incubating	Interface REST para interagir com o Apache Spark
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	0.12.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.0.176	Drivers NVIDIA e toolkit CUDA
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.

Componente	Versão	Descrição
phoenix-library	4.11.0- -1.3 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.11.0- -1.3 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.187	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.187	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
spark-client	2.2.1	Os clientes da linha de comando do Spark.
spark-history-server	2.2.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.2.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.2.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.

Componente	Versão	Descrição
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.3	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.11.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.11.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.

Classificações	Descrição
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.

Classificações	Descrição
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.

Classificações	Descrição
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.

Classificações	Descrição
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.

Classificações	Descrição
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.

Classificações	Descrição
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.

Classificações	Descrição
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.10.1 do Amazon EMR

Versões 5.10.1 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, e ZooKeeper.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.10.1	emr-5.10.0	emr-5.9.1	emr-5.9.0
AWS SDK for Java	1.11.21	1.11.21	1.11.183	1.11.183
Python	2.7, 3.4	2.7, 3.4	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-5.10.1	emr-5.10.0	emr-5.9.1	emr-5.9.0
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.1	2.3.1	2.3.0	2.3.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.1	2.3.1	2.3.0	2.3.0
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	0.12.0	0.12.0	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.184	0.184
Spark	2.2.0	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6

	emr-5.10.1	emr-5.10.0	emr-5.9.1	emr-5.9.0
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.10.1

Esta é uma versão de patch para adicionar a autenticação AWS Signature versão 4 para solicitações ao Amazon S3. Todas as aplicações e componentes são os mesmos que os da versão anterior do Amazon EMR.

> **Important**
>
> Nesta versão de lançamento, o Amazon EMR usa o AWS Signature versão 4 exclusivamente para autenticar solicitações para o Amazon S3. Para obter mais informações, consulte Novidades.

Versões de componente 5.10.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado

`myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.5.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.7.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.20.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.3.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-5	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-5	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-5	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-5	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-5	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-5	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.7.3-amzn-5	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-5	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.7.3-amzn-5	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-5	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.3.1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.3.1	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.3.1	HBase cliente de linha de comando.
hbase-rest-server	1.3.1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.3.1	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.1-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.1-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	2.3.1-amzn-0	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.1-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.3.1-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.3.1-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.1-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.0.1	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
livy-server	0.4.0-incubating	Interface REST para interagir com o Apache Spark
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	0.12.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.0.176	Drivers NVIDIA e toolkit CUDA

Componente	Versão	Descrição
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.11.0- -1.3 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.11.0- -1.3 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.187	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.187	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
spark-client	2.2.0	Os clientes da linha de comando do Spark.
spark-history-server	2.2.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.

Componente	Versão	Descrição
spark-on-yarn	2.2.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.2.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.3	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.10.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.10.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.

Classificações	Descrição
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive- site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.pro perties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.pr operties do Hive.

Classificações	Descrição
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.proper ties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.x ml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properti es do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.

Classificações	Descrição
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.

Classificações	Descrição
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.

Classificações	Descrição
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.

Classificações	Descrição
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.10.0 do Amazon EMR

Versões 5.10.0 da aplicação

Essa versão inclui os seguintes aplicativos: Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, e ZooKeeper.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.10.0	emr-5.9.1	emr-5.9.0	emr-5.8.3
AWS SDK for Java	1.11.21	1.11.183	1.11.183	1.11.160

	emr-5.10.0	emr-5.9.1	emr-5.9.0	emr-5.8.3
Python	2.7, 3.4	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.1	2.3.0	2.3.0	2.3.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.1	2.3.0	2.3.0	2.3.0
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	-
MXNet	0.12.0	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0

	emr-5.10.0	emr-5.9.1	emr-5.9.0	emr-5.8.3
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.16.0
Presto	0.187	0.184	0.184	0.170
Spark	2.2.0	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.10.0

As notas da versão a seguir incluem informações para a versão 5.10.0 do Amazon EMR. As alterações são referentes à versão 5.9.0 do Amazon EMR.

Atualizações

- AWS SDK para Java 1.11.221
- Hive 2.3.1
- Presto 0.187

Novos recursos

- Adicionado o suporte para autenticação do Kerberos. Para obter mais informações, consulte Usar autenticação Kerberos no Guia de gerenciamento do Amazon EMR.

- Adicionado suporte para perfis do IAM para EMRFS. Para obter mais informações, consulte Configure IAM Roles for EMRFS Requests to Amazon S3 no Guia de gerenciamento do Amazon EMR.
- Suporte adicionado para tipos de instância P2 e P3 baseados em GPU. Para obter mais informações, consulte Instâncias EC2 P2 da Amazon e Instâncias EC2 P3 da Amazon. Os drivers NVIDIA 384.81 e CUDA driver 9.0.176 são instalados nesses tipos de instância por padrão.
- O suporte adicionado para Apache MXNet.

Alterações, melhorias e problemas resolvidos

- Presto
  - Foi adicionado suporte para usar o AWS Glue Data Catalog como metastore padrão do Hive. Para obter mais informações, consulte Usando o Presto com o AWS Glue Data Catalog.
  - Suporte adicionado para funções geoespaciais.
  - Suporte de vazamento para disco adicionado para uniões.
  - Suporte adicionado para o Conector Redshift.
- Spark
  - Backport SPARK-20640, que causa tempo limite rpc e as repetições de valores de registro configuráveis usando as propriedades `spark.shuffle.registration.timeout` e `spark.shuffle.registration.maxAttempts`.
  - SPARK-21549 retroportado, que corrige um erro que ocorre ao gravar de forma personalizada em locais não HDFS. OutputFormat
- Backport Hadoop-13270
- As bibliotecas Numpy, Scipy e Matplotlib foram removidas da AMI base do Amazon EMR. Se forem necessárias para o aplicativo, essas bibliotecas estarão disponíveis no repositório do aplicativo. Portanto, você pode usar uma ação de bootstrap para instalá-las em todos os nós usando `yum install`.
- A AMI base do Amazon EMR não tem mais pacotes do RPM de aplicações incluídos, de maneira que os pacotes do RPM não estão mais presentes em nós de cluster. A AMI personalizada AMIs e a AMI básica do Amazon EMR agora fazem referência ao repositório de pacotes RPM no Amazon S3.
- Devido à introdução do faturamento por segundo na Amazon EC2, o comportamento padrão de redução de escala agora é Encerrar na conclusão da tarefa, em vez de Encerrar na hora

da instância. Para obter mais informações, consulte Configure Cluster Scale-Down (Configurar escalabilidade vertical do cluster).

Problemas conhecidos

- MXNet não inclui bibliotecas OpenCV.
- Por padrão, o Hive 2.3.1 define `hive.compute.query.using.stats=true`. Isso faz com que as consultas obtenham dados de estatísticas existentes em vez de diretamente dos dados, o que pode gerar confusão. Por exemplo, se você tiver uma tabela com `hive.compute.query.using.stats=true` e fizer upload de novos arquivos para a tabela `LOCATION`, a execução de uma consulta `SELECT COUNT(*)` na tabela retornará a contagem das estatísticas, e não selecionará as linhas adicionadas.

  Como alternativa, use o comando `ANALYZE TABLE` para reunir novas estatísticas ou defina `hive.compute.query.using.stats=false`. Para obter mais informações, consulte Estatísticas no Hive na documentação do Apache Hive.

Versões de componente 5.10.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.5.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.

Componente	Versão	Descrição
emr-goodies	2.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.7.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.20.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.3.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-5	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.

Componente	Versão	Descrição
hadoop-hdfs-datanode	2.7.3-amzn-5	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-5	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-5	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-5	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-5	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-5	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-5	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.3-amzn-5	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-5	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.

Componente	Versão	Descrição
hbase-hmaster	1.3.1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.3.1	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.3.1	HBase cliente de linha de comando.
hbase-rest-server	1.3.1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.3.1	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.1-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.1-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.3.1-amzn-0	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.1-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.3.1-amzn-0	Cliente do Hive-hbase.

Componente	Versão	Descrição
hive-metastore-server	2.3.1-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.1-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.0.1	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
livy-server	0.4.0-incubating	Interface REST para interagir com o Apache Spark
mahout-client	0.13.0	A biblioteca de Machine Learning.
mxnet	0.12.0	Uma biblioteca flexível, escalável e eficiente para aprendizado aprofundado.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
nvidia-cuda	9.0.176	Drivers NVIDIA e toolkit CUDA
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.

Componente	Versão	Descrição
phoenix-library	4.11.0- -1.3 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.11.0- -1.3 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.187	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.187	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
spark-client	2.2.0	Os clientes da linha de comando do Spark.
spark-history-server	2.2.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.2.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.2.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.

Componente	Versão	Descrição
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.3	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.10.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.10.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.

Classificações	Descrição
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.

Classificações	Descrição
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.

Classificações	Descrição
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.

Classificações	Descrição
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.

Classificações	Descrição
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.

Classificações	Descrição
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.

Classificações	Descrição
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.9.1 do Amazon EMR

Versões 5.9.1 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGanglia, HBase, HCatalog, HadoopHive, Hue, Livy, Mahout, Oozie, Phoenix, Pig, Presto, SparkSqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.9.1	emr-5.9.0	emr-5.8.3	emr-5.8.2
AWS SDK for Java	1.11.183	1.11.183	1.11.160	1.11.160
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-5.9.1	emr-5.9.0	emr-5.8.3	emr-5.8.2
Flink	1.3.2	1.3.2	1.3.1	1.3.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.0	2.3.0	2.3.0	2.3.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.3.0	2.3.0	2.3.0
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.16.0	0.16.0
Presto	0.184	0.184	0.170	0.170
Spark	2.2.0	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6

	emr-5.9.1	emr-5.9.0	emr-5.8.3	emr-5.8.2
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.9.1

Esta é uma versão de patch para adicionar a autenticação AWS Signature versão 4 para solicitações ao Amazon S3. Todas as aplicações e componentes são os mesmos que os da versão anterior do Amazon EMR.

> **Important**
>
> Nesta versão de lançamento, o Amazon EMR usa o AWS Signature versão 4 exclusivamente para autenticar solicitações para o Amazon S3. Para obter mais informações, consulte Novidades.

Versões de componente 5.9.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado

`myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.4.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.7.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.19.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.3.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-4	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-4	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-4	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-4	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-4	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-4	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.7.3-amzn-4	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-4	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.7.3-amzn-4	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-4	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.3.1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.3.1	Serviço para atender a uma ou mais HBase regiões.
hbase-client	1.3.1	HBase cliente de linha de comando.
hbase-rest-server	1.3.1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.3.1	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.0-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.0-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	2.3.0-amzn-0	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.0-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.3.0-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.3.0-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.0-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.0.1	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
livy-server	0.4.0-incubating	Interface REST para interagir com o Apache Spark
mahout-client	0.13.0	A biblioteca de Machine Learning.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
oozie-client	4.3.0	O cliente da linha de comando do Oozie.

Componente	Versão	Descrição
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.11.0- -1.3 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.11.0- -1.3 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.184	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.184	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
spark-client	2.2.0	Os clientes da linha de comando do Spark.
spark-history-server	2.2.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.2.0	O mecanismo de execução em memória para o YARN.

Componente	Versão	Descrição
spark-yarn-slave	2.2.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.2	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.9.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.9.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.

Classificações	Descrição
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive- site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.pro perties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.pr operties do Hive.

Classificações	Descrição
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.proper ties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.x ml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properti es do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.

Classificações	Descrição
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.

Classificações	Descrição
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.

Classificações	Descrição
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.

Classificações	Descrição
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.9.0 do Amazon EMR

Versões 5.9.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGanglia, HBase, HCatalog, HadoopHive, Hue, Livy, Mahout, Oozie, Phoenix, Pig, Presto, SparkSqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.9.0	emr-5.8.3	emr-5.8.2	emr-5.8.1
AWS SDK for Java	1.11.183	1.11.160	1.11.160	1.11.160

	emr-5.9.0	emr-5.8.3	emr-5.8.2	emr-5.8.1
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.1	1.3.1	1.3.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.0	2.3.0	2.3.0	2.3.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.3.0	2.3.0	2.3.0
Hudi	-	-	-	-
Hue	4.0.1	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0

	emr-5.9.0	emr-5.8.3	emr-5.8.2	emr-5.8.1
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.16.0	0.16.0	0.16.0
Presto	0.184	0.170	0.170	0.170
Spark	2.2.0	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.9.0

As notas da versão a seguir incluem informações para a versão 5.9.0 do Amazon EMR. As alterações são referentes à versão 5.8.0 do Amazon EMR.

Data do release: 5 de outubro de 2017

Última atualização de recursos: 12 de outubro de 2017

Atualizações

- AWS SDK para Java versão 1.11.183
- Flink 1.3.2
- Hue 4.0.1
- Pig 0.17.0
- Presto 0.184

Novos recursos

- Adição do suporte ao Livy (versão Livy 0.4.0 - em incubação). Para obter mais informações, consulte Apache Livy.
- Adição de suporte para Hue Notebook para Spark.
- Foi adicionado suporte para EC2 instâncias Amazon da série i3 (12 de outubro de 2017).

Alterações, melhorias e problemas resolvidos

- Spark
  - Adição de um novo conjunto de recursos que ajudam a garantir que o Spark lide de uma forma mais fácil com o encerramento de nós devido a um redimensionamento manual ou uma solicitação de política de escalabilidade automática. Para obter mais informações, consulte Configurar o comportamento de desativação de nós.
  - O SSL é usado em vez do 3DES para criptografia em trânsito para o serviço de transferência em bloco, o que melhora o desempenho ao usar tipos de EC2 instância da Amazon com AES-NI.
  - SPARK-21494 enviado para backport.
- Zeppelin
  - ZEPPELIN-2377 enviado para backport.
- HBase
  - Foi adicionado o patch HBASE-18533, que permite valores adicionais para HBase BucketCache configuração usando a classificação de configuração. `hbase-site`
- Hue
  - Foi adicionado suporte ao AWS Glue Data Catalog para o editor de consultas Hive no Hue.
  - Por padrão, os superusuários do Hue podem acessar todos os arquivos que os perfis do IAM do Amazon EMR têm permissão para acessar. Os usuários recém-criados não têm automaticamente permissões para acessar o navegador de arquivos Amazon S3 e devem ter as permissões `filebrowser.s3_access` ativadas para o grupo deles.
- Resolveu um problema que fazia com que os dados JSON subjacentes criados usando o AWS Glue Data Catalog ficassem inacessíveis.

Problemas conhecidos

- A inicialização do cluster falha quando todas as aplicações são instaladas e o tamanho padrão do volume raiz do Amazon EBS não é alterado. Como solução alternativa, use o `aws emr create-cluster` comando do AWS CLI e especifique um `--ebs-root-volume-size` parâmetro maior.
- Por padrão, o Hive 2.3.0 define `hive.compute.query.using.stats=true`. Isso faz com que as consultas obtenham dados de estatísticas existentes em vez de diretamente dos dados, o que pode gerar confusão. Por exemplo, se você tiver uma tabela com `hive.compute.query.using.stats=true` e fizer upload de novos arquivos para a tabela `LOCATION`, a execução de uma consulta `SELECT COUNT(*)` na tabela retornará a contagem das estatísticas, e não selecionará as linhas adicionadas.

  Como alternativa, use o comando `ANALYZE TABLE` para reunir novas estatísticas ou defina `hive.compute.query.using.stats=false`. Para obter mais informações, consulte Estatísticas no Hive na documentação do Apache Hive.

Versões de componente 5.9.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.4.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.

Componente	Versão	Descrição
emr-goodies	2.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.7.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.19.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.3.2	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-4	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.

Componente	Versão	Descrição
hadoop-hdfs-datanode	2.7.3-amzn-4	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-4	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-4	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-4	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-4	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-4	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-4	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.3-amzn-4	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-4	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.

Componente	Versão	Descrição
hbase-hmaster	1.3.1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.3.1	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.3.1	HBase cliente de linha de comando.
hbase-rest-server	1.3.1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.3.1	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.0-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.0-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.3.0-amzn-0	Endpoint HTTP fornecend o uma interface REST para HCatalog.
hive-client	2.3.0-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.3.0-amzn-0	Cliente do Hive-hbase.

Componente	Versão	Descrição
hive-metastore-server	2.3.0-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.0-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	4.0.1	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
livy-server	0.4.0-incubating	Interface REST para interagir com o Apache Spark
mahout-client	0.13.0	A biblioteca de Machine Learning.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.11.0- -1.3 HBase	As bibliotecas do phoenix para servidor e cliente

Componente	Versão	Descrição
phoenix-query-server	4.11.0- -1.3 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.184	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.184	O serviço que executa partes de uma consulta.
pig-client	0.17.0	O cliente da linha de comando do Pig.
spark-client	2.2.0	Os clientes da linha de comando do Spark.
spark-history-server	2.2.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.2.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.2.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.

Componente	Versão	Descrição
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.2	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.9.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.9.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.

Classificações	Descrição
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.

Classificações	Descrição
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.

Classificações	Descrição
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
livy-conf	Alterar os valores no arquivo livy.conf do Livy.
livy-env	Alterar os valores no ambiente do Livy.
livy-log4j	Alterar as configurações de log4j.properties no Livy.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.

Classificações	Descrição
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.

Classificações	Descrição
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.

Classificações	Descrição
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.

Classificações	Descrição
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.8.3 do Amazon EMR

Versões 5.8.3 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGanglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, SparkSqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.8.3	emr-5.8.2	emr-5.8.1	emr-5.8.0
AWS SDK for Java	1.11.160	1.11.160	1.11.160	1.11.160
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-5.8.3	emr-5.8.2	emr-5.8.1	emr-5.8.0
Flink	1.3.1	1.3.1	1.3.1	1.3.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.0	2.3.0	2.3.0	2.3.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.3.0	2.3.0	2.3.0
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.2.0	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6

	emr-5.8.3	emr-5.8.2	emr-5.8.1	emr-5.8.0
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.8.3

Esta é uma versão de patch para adicionar a autenticação AWS Signature versão 4 para solicitações ao Amazon S3. Todas as aplicações e componentes são os mesmos que os da versão anterior do Amazon EMR.

> **Important**
>
> Nesta versão de lançamento, o Amazon EMR usa o AWS Signature versão 4 exclusivamente para autenticar solicitações para o Amazon S3. Para obter mais informações, consulte Novidades.

Versões de componente 5.8.3

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado

`myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.4.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.6.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.18.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.3.1	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-3	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-3	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-3	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-3	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-3	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-3	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.7.3-amzn-3	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-3	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.7.3-amzn-3	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-3	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.3.1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.3.1	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.3.1	HBase cliente de linha de comando.
hbase-rest-server	1.3.1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.3.1	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.0-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.0-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	2.3.0-amzn-0	Endpoint HTTP fornecend o uma interface REST para HCatalog.
hive-client	2.3.0-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.3.0-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.3.0-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.0-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.12.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.13.0	A biblioteca de Machine Learning.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.

Componente	Versão	Descrição
phoenix-library	4.11.0- -1.3 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.11.0- -1.3 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.170	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.170	O serviço que executa partes de uma consulta.
pig-client	0.16.0-amzn-1	O cliente da linha de comando do Pig.
spark-client	2.2.0	Os clientes da linha de comando do Spark.
spark-history-server	2.2.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.2.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.2.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.

Componente	Versão	Descrição
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.2	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.8.3

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.8.3

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.

Classificações	Descrição
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.

Classificações	Descrição
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.

Classificações	Descrição
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.

Classificações	Descrição
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.

Classificações	Descrição
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.

Classificações	Descrição
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.8.2 do Amazon EMR

Versões 5.8.2 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGanglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, SparkSqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.8.2	emr-5.8.1	emr-5.8.0	emr-5.7.1
AWS SDK for Java	1.11.160	1.11.160	1.11.160	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.1	1.3.1	1.3.1	1.3.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1

	emr-5.8.2	emr-5.8.1	emr-5.8.0	emr-5.7.1
HCatalog	2.3.0	2.3.0	2.3.0	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.3.0	2.3.0	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.2.0	2.2.0	2.2.0	2.1.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.8.2	emr-5.8.1	emr-5.8.0	emr-5.7.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.8.2

As notas da versão a seguir incluem informações para a versão 5.8.2 do Amazon EMR. As alterações são referentes à versão 5.8.1.

Data da versão inicial: 29 de março de 2018

Alterações, melhorias e problemas resolvidos

- Atualizado o kernel do Amazon Linux da AMI padrão do Amazon Linux para Amazon EMR para abordar possíveis vulnerabilidades.

Versões de componente 5.8.2

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.4.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.6.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.18.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.3.1	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas

Componente	Versão	Descrição
		coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-3	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-3	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-3	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-3	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-3	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-3	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-3	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-3	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.7.3-amzn-3	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-3	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.3.1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.3.1	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.3.1	HBase cliente de linha de comando.
hbase-rest-server	1.3.1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.3.1	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.0-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.0-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	2.3.0-amzn-0	Endpoint HTTP fornecend o uma interface REST para HCatalog.
hive-client	2.3.0-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.3.0-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.3.0-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.0-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.12.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.13.0	A biblioteca de Machine Learning.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.

Componente	Versão	Descrição
phoenix-library	4.11.0- -1.3 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.11.0- -1.3 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.170	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.170	O serviço que executa partes de uma consulta.
pig-client	0.16.0-amzn-1	O cliente da linha de comando do Pig.
spark-client	2.2.0	Os clientes da linha de comando do Spark.
spark-history-server	2.2.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.2.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.2.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.

Componente	Versão	Descrição
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.2	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.8.2

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.8.2

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.

Classificações	Descrição
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.

Classificações	Descrição
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.

Classificações	Descrição
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.

Classificações	Descrição
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.

Classificações	Descrição
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.

Classificações	Descrição
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.8.1 do Amazon EMR

Versões 5.8.1 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGanglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, SparkSqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.8.1	emr-5.8.0	emr-5.7.1	emr-5.7.0
AWS SDK for Java	1.11.160	1.11.160	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.1	1.3.1	1.3.0	1.3.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1

	emr-5.8.1	emr-5.8.0	emr-5.7.1	emr-5.7.0
HCatalog	2.3.0	2.3.0	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.3.0	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.2.0	2.2.0	2.1.1	2.1.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.8.1	emr-5.8.0	emr-5.7.1	emr-5.7.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.8.1

As notas da versão a seguir incluem informações para a versão 5.8.1 do Amazon EMR. As alterações são referentes à versão 5.8.0 do Amazon EMR.

Data da versão inicial: 22 de janeiro de 2018

Alterações, melhorias e problemas resolvidos

- Atualizado o kernel do Amazon Linux da AMI padrão do Amazon Linux para Amazon EMR para abordar vulnerabilidades associadas à execução especulativa (CVE-2017-5715, CVE-2017-5753 e CVE-2017-5754). Para obter mais informações, consulte https://aws.amazon.com/security/security-bulletins/AWS-2018-013/.

Versões de componente 5.8.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.4.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.6.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.18.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.3.1	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas

Componente	Versão	Descrição
		coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-3	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-3	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-3	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-3	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-3	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-3	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-3	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-3	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.7.3-amzn-3	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-3	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.3.1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.3.1	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.3.1	HBase cliente de linha de comando.
hbase-rest-server	1.3.1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.3.1	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.0-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.0-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	2.3.0-amzn-0	Endpoint HTTP fornecend o uma interface REST para HCatalog.
hive-client	2.3.0-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.3.0-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.3.0-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.0-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.12.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.13.0	A biblioteca de Machine Learning.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.

Componente	Versão	Descrição
phoenix-library	4.11.0- -1.3 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.11.0- -1.3 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.170	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.170	O serviço que executa partes de uma consulta.
pig-client	0.16.0-amzn-1	O cliente da linha de comando do Pig.
spark-client	2.2.0	Os clientes da linha de comando do Spark.
spark-history-server	2.2.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.2.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.2.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.

Componente	Versão	Descrição
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.2	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.8.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.8.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.

Classificações	Descrição
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.

Classificações	Descrição
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.

Classificações	Descrição
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.

Classificações	Descrição
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.

Classificações	Descrição
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.

Classificações	Descrição
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.8.0 do Amazon EMR

Versões 5.8.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGanglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, SparkSqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.8.0	emr-5.7.1	emr-5.7.0	emr-5.6.1
AWS SDK for Java	1.11.160	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.1	1.3.0	1.3.0	1.2.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.0

	emr-5.8.0	emr-5.7.1	emr-5.7.0	emr-5.6.1
HCatalog	2.3.0	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.9.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.2.0	2.1.1	2.1.1	2.1.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.8.0	emr-5.7.1	emr-5.7.0	emr-5.6.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.8.0

As notas da versão a seguir incluem informações para a versão 5.8.0 do Amazon EMR. As alterações são referentes à versão 5.7.0 do Amazon EMR.

Data da versão inicial: 10 de agosto de 2017

Última atualização de recurso: 25 de setembro de 2017

Atualizações

- AWS SDK 1.11.160
- Flink 1.3.1
- Hive 2.3.0. Para obter mais informações, consulte Notas de release no site do Apache Hive.
- Spark 2.2.0. Para obter mais informações, consulte Notas de release no site do Apache Spark.

Novos recursos

- Adição de suporte para visualização do histórico de aplicativos (25 de setembro de 2017). Para obter mais informações, consulte Visualizar histórico de aplicações no Guia de gerenciamento do Amazon EMR.

Alterações, melhorias e problemas resolvidos

- Integração com o AWS Glue Data Catalog
  - Foi adicionada a capacidade do Hive e do Spark SQL de usar o AWS Glue Data Catalog como armazenamento de metadados do Hive. Para obter mais informações, consulte Usar o Catálogo de Dados do AWS Glue como metastore do Hive e Use o catálogo do AWS Glue Data Catalog com o Spark no Amazon EMR.

- Adicionado o Application history (Histórico do aplicativo) aos detalhes do cluster, o que permite que você visualize dados históricos de aplicativos do YARN e detalhes adicionais de aplicativos do Spark. Para obter mais informações, consulte Visualizar histórico de aplicações no Guia de gerenciamento do Amazon EMR.
- Oozie
  - OOZIE-2748 enviado para backport.
- Hue
  - HUE-5859 enviado para backport
- HBase
  - Foi adicionado um patch para expor a hora de início do servidor HBase mestre por meio do Java Management Extensions (JMX) usando. `getMasterInitializedTime`
  - Adicionado um patch que melhora a hora de início do cluster.

Problemas conhecidos

- A inicialização do cluster falha quando todas as aplicações são instaladas e o tamanho padrão do volume raiz do Amazon EBS não é alterado. Como solução alternativa, use o `aws emr create-cluster` comando do AWS CLI e especifique um `--ebs-root-volume-size` parâmetro maior.
- Por padrão, o Hive 2.3.0 define `hive.compute.query.using.stats=true`. Isso faz com que as consultas obtenham dados de estatísticas existentes em vez de diretamente dos dados, o que pode gerar confusão. Por exemplo, se você tiver uma tabela com `hive.compute.query.using.stats=true` e fizer upload de novos arquivos para a tabela `LOCATION`, a execução de uma consulta `SELECT COUNT(*)` na tabela retornará a contagem das estatísticas, e não selecionará as linhas adicionadas.

  Como alternativa, use o comando `ANALYZE TABLE` para reunir novas estatísticas ou defina `hive.compute.query.using.stats=false`. Para obter mais informações, consulte Estatísticas no Hive na documentação do Apache Hive.
- Spark: ao usar o Spark, há um problema de vazamento no manipulador de arquivos com o daemon apppusher, o que pode ocorrer para um trabalho de execução prolongada do Spark depois de várias horas ou dias. Para corrigir o problema, conecte-se ao nó primário e digite `sudo /etc/init.d/apppusher stop`. Isso interrompe o daemon apppusher, que o Amazon EMR reiniciará automaticamente.
- Application history
  - Os dados históricos dos executores inativos do Spark não está disponível.

- O histórico do aplicativo não está disponível para clusters que usam uma configuração de segurança para habilitar a criptografia em andamento.

Versões de componente 5.8.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.4.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.4.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.4.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.6.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.18.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.

Componente	Versão	Descrição
flink-client	1.3.1	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-3	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-3	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-3	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-3	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-3	O endpoint de HTTP para as operações do HDFS.

Componente	Versão	Descrição
hadoop-kms-server	2.7.3-amzn-3	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-3	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-3	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.3-amzn-3	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-3	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.3.1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.3.1	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.3.1	HBase cliente de linha de comando.

Componente	Versão	Descrição
hbase-rest-server	1.3.1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.3.1	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.3.0-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.3.0-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.3.0-amzn-0	Endpoint HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.3.0-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.3.0-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.3.0-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.3.0-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.

Componente	Versão	Descrição
hue-server	3.12.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.13.0	A biblioteca de Machine Learning.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.11.0- -1.3 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.11.0- -1.3 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.170	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.170	O serviço que executa partes de uma consulta.
pig-client	0.16.0-amzn-1	O cliente da linha de comando do Pig.

Componente	Versão	Descrição
spark-client	2.2.0	Os clientes da linha de comando do Spark.
spark-history-server	2.2.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.2.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.2.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.2	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.8.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.8.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop

Classificações	Descrição
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive- site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.

Classificações	Descrição
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.

Classificações	Descrição
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.

Classificações	Descrição
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.

Classificações	Descrição
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.

Classificações	Descrição
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.7.1 do Amazon EMR

Versões 5.7.1 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGanglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, SparkSqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.7.1	emr-5.7.0	emr-5.6.1	emr-5.6.0
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8

	emr-5.7.1	emr-5.7.0	emr-5.6.1	emr-5.6.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.0	1.3.0	1.2.1	1.2.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.9.0	4.9.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0

	emr-5.7.1	emr-5.7.0	emr-5.6.1	emr-5.6.0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.1	2.1.1	2.1.1	2.1.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.7.1

Esta é uma versão de patch para adicionar a autenticação AWS Signature versão 4 para solicitações ao Amazon S3. Todas as aplicações e componentes são os mesmos que os da versão anterior do Amazon EMR.

> **Important**
>
> Nesta versão de lançamento, o Amazon EMR usa o AWS Signature versão 4 exclusivamente para autenticar solicitações para o Amazon S3. Para obter mais informações, consulte Novidades.

Versões de componente 5.7.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a

versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.3.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.3.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.3.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.5.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.18.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.3.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-2	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-2	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-2	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-2	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-2	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.

Componente	Versão	Descrição
hadoop-yarn-nodemanager	2.7.3-amzn-2	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hadoop-yarn-timeline-server	2.7.3-amzn-2	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.3.1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.3.1	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.3.1	HBase cliente de linha de comando.
hbase-rest-server	1.3.1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.3.1	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.1.1-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.

Componente	Versão	Descrição
hcatalog-server	2.1.1-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.1.1-amzn-0	Endpoint HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.1.1-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.1.1-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.1.1-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.1.1-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.12.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.13.0	A biblioteca de Machine Learning.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
oozie-client	4.3.0	O cliente da linha de comando do Oozie.

Componente	Versão	Descrição
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.11.0- -1.3 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.11.0- -1.3 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.170	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.170	O serviço que executa partes de uma consulta.
pig-client	0.16.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	2.1.1	Os clientes da linha de comando do Spark.
spark-history-server	2.1.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.1.1	O mecanismo de execução em memória para o YARN.

Componente	Versão	Descrição
spark-yarn-slave	2.1.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.2	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.7.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.7.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.

Classificações	Descrição
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive- site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.pro perties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.pr operties do Hive.

Classificações	Descrição
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.proper ties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.x ml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properti es do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.

Classificações	Descrição
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.

Classificações	Descrição
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.

Classificações	Descrição
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.

Classificações	Descrição
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.7.0 do Amazon EMR

Versões 5.7.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGanglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, SparkSqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.7.0	emr-5.6.1	emr-5.6.0	emr-5.5.4
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-5.7.0	emr-5.6.1	emr-5.6.0	emr-5.5.4
Flink	1.3.0	1.2.1	1.2.1	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.9.0	4.9.0	4.9.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.1	2.1.1	2.1.1	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6

	emr-5.7.0	emr-5.6.1	emr-5.6.0	emr-5.5.4
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.1	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.7.0

As notas da versão a seguir incluem informações para a versão 5.7.0 do Amazon EMR. As alterações são referentes à versão 5.6.0 do Amazon EMR.

Data do release: 13 de julho de 2017

Atualizações

- Flink 1.3.0
- Phoenix 4.11.0
- Zeppelin 0.7.2

Novos recursos

- Adicionada a capacidade de especificação de uma AMI do Amazon Linux personalizada ao ser criado um cluster. Para obter mais informações, consulte Usando uma AMI personalizada.

Alterações, melhorias e problemas resolvidos

- HBase
  - Capacidade adicional para configurar clusters de HBase réplica de leitura. Consulte Usando um cluster de réplica de leitura.
  - Várias correções de erros e melhorias
- Presto: adicionada a capacidade de configurar `node.properties`.

- YARN :adicionada a capacidade de configurar `container-log4j.properties`
- Sqoop: enviado para backport SQOOP-2880, que apresenta um argumento que permite definir o diretório temporário do Sqoop.

Versões de componente 5.7.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.3.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.3.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.3.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.5.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.18.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.

Componente	Versão	Descrição
flink-client	1.3.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-2	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-2	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-2	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-2	O endpoint de HTTP para as operações do HDFS.

Componente	Versão	Descrição
hadoop-kms-server	2.7.3-amzn-2	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-2	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.3-amzn-2	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-2	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.3.1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.3.1	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.3.1	HBase cliente de linha de comando.

Componente	Versão	Descrição
hbase-rest-server	1.3.1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.3.1	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.1.1-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.1.1-amzn-0	Endpoint HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.1.1-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.1.1-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.1.1-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.1.1-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.

Componente	Versão	Descrição
hue-server	3.12.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.13.0	A biblioteca de Machine Learning.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.11.0- -1.3 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.11.0- -1.3 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.170	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.170	O serviço que executa partes de uma consulta.
pig-client	0.16.0-amzn-0	O cliente da linha de comando do Pig.

Componente	Versão	Descrição
spark-client	2.1.1	Os clientes da linha de comando do Spark.
spark-history-server	2.1.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.1.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.1.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.2	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.7.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.7.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop

Classificações	Descrição
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive- site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.

Classificações	Descrição
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.

Classificações	Descrição
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.

Classificações	Descrição
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.

Classificações	Descrição
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.

Classificações	Descrição
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.6.1 do Amazon EMR

Versões 5.6.1 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGanglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, SparkSqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.6.1	emr-5.6.0	emr-5.5.4	emr-5.5.3
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8

	emr-5.6.1	emr-5.6.0	emr-5.5.4	emr-5.5.3
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.1	1.2.1	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0

	emr-5.6.1	emr-5.6.0	emr-5.5.4	emr-5.5.3
Presto	0.170	0.170	0.170	0.170
Spark	2.1.1	2.1.1	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.6.1

Esta é uma versão de patch para adicionar a autenticação AWS Signature versão 4 para solicitações ao Amazon S3. Todas as aplicações e componentes são os mesmos que os da versão anterior do Amazon EMR.

> **Important**
>
> Nesta versão de lançamento, o Amazon EMR usa o AWS Signature versão 4 exclusivamente para autenticar solicitações para o Amazon S3. Para obter mais informações, consulte Novidades.

Versões de componente 5.6.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a

versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *CommunityVersion*-amzn-*EmrVersion*. O *EmrVersion* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.3.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.3.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.3.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.5.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.17.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.2.1	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-2	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-2	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-2	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-2	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-2	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.

Componente	Versão	Descrição
hadoop-yarn-nodemanager	2.7.3-amzn-2	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hadoop-yarn-timeline-server	2.7.3-amzn-2	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.3.1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.3.1	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.3.1	HBase cliente de linha de comando.
hbase-rest-server	1.3.1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.3.1	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.1.1-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.

Componente	Versão	Descrição
hcatalog-server	2.1.1-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.1.1-amzn-0	Endpoint HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.1.1-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.1.1-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.1.1-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.1.1-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.12.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.13.0	A biblioteca de Machine Learning.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
oozie-client	4.3.0	O cliente da linha de comando do Oozie.

Componente	Versão	Descrição
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.9.0- -1.2 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.9.0- -1.2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.170	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.170	O serviço que executa partes de uma consulta.
pig-client	0.16.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	2.1.1	Os clientes da linha de comando do Spark.
spark-history-server	2.1.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.1.1	O mecanismo de execução em memória para o YARN.

Componente	Versão	Descrição
spark-yarn-slave	2.1.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.6.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.6.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.

Classificações	Descrição
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive- site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.pro perties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.pr operties do Hive.

Classificações	Descrição
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.proper ties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.x ml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properti es do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.

Classificações	Descrição
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.

Classificações	Descrição
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.

Classificações	Descrição
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.

Classificações	Descrição
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.6.0 do Amazon EMR

Versões 5.6.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGanglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, SparkSqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.6.0	emr-5.5.4	emr-5.5.3	emr-5.5.2
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-5.6.0	emr-5.5.4	emr-5.5.3	emr-5.5.2
Flink	1.2.1	1.2.0	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.1	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6

	emr-5.6.0	emr-5.5.4	emr-5.5.3	emr-5.5.2
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.6.0

As notas da versão a seguir incluem informações para a versão 5.6.0 do Amazon EMR. As alterações são referentes à versão 5.5.0 do Amazon EMR.

Data do release: 5 de junho de 2017

Atualizações

- Flink 1.2.1
- HBase 1.3.1
- Mahout 0.13.0. Esta é a primeira versão do Mahout compatível com o Spark 2.x nas versões 5.0 e posteriores do Amazon EMR.
- Spark 2.1.1

Alterações, melhorias e problemas resolvidos

- Presto
  - Foi adicionada a capacidade de permitir a comunicação SSL/TLS segura entre os nós do Presto, ativando a criptografia em trânsito usando uma configuração de segurança. Para obter mais informações, consulte Criptografia de dados em trânsito.
  - Enviado para backport o Presto 7661, o que adiciona a opção `VERBOSE` à instrução `EXPLAIN ANALYZE` para relatar estatísticas de baixo nível mais detalhadas sobre um plano de consulta.

Versões de componente 5.6.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.3.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.3.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.3.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.5.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.17.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.2.1	Scripts do cliente da linha de comando e aplicativos do Apache Flink.

Componente	Versão	Descrição
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-2	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-2	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-2	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-2	O endpoint de HTTP para as operações do HDFS.

Componente	Versão	Descrição
hadoop-kms-server	2.7.3-amzn-2	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-2	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.3-amzn-2	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-2	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.3.1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.3.1	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.3.1	HBase cliente de linha de comando.

Componente	Versão	Descrição
hbase-rest-server	1.3.1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.3.1	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.1.1-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.1.1-amzn-0	Endpoint HTTP fornecend o uma interface REST para HCatalog.
hive-client	2.1.1-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.1.1-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.1.1-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.1.1-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.

Componente	Versão	Descrição
hue-server	3.12.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.13.0	A biblioteca de Machine Learning.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.9.0- -1.2 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.9.0- -1.2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.170	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.170	O serviço que executa partes de uma consulta.
pig-client	0.16.0-amzn-0	O cliente da linha de comando do Pig.

Componente	Versão	Descrição
spark-client	2.1.1	Os clientes da linha de comando do Spark.
spark-history-server	2.1.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.1.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.1.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.6.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.6.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop

Classificações	Descrição
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive- site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.

Classificações	Descrição
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.

Classificações	Descrição
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.

Classificações	Descrição
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-env	Altere os valores no arquivo presto-env.sh do Presto.
presto-node	Altere os valores no arquivo node.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.

Classificações	Descrição
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.

Classificações	Descrição
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.5.4 do Amazon EMR

Versões 5.5.4 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGanglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, SparkSqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.5.4	emr-5.5.3	emr-5.5.2	emr-5.5.1
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8

	emr-5.5.4	emr-5.5.3	emr-5.5.2	emr-5.5.1
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.2.0	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0

	emr-5.5.4	emr-5.5.3	emr-5.5.2	emr-5.5.1
Presto	0.170	0.170	0.170	0.170
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.5.4

Esta é uma versão de patch para adicionar a autenticação AWS Signature versão 4 para solicitações ao Amazon S3. Todas as aplicações e componentes são os mesmos que os da versão anterior do Amazon EMR.

> **Important**
>
> Nesta versão de lançamento, o Amazon EMR usa o AWS Signature versão 4 exclusivamente para autenticar solicitações para o Amazon S3. Para obter mais informações, consulte Novidades.

Versões de componente 5.5.4

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a

versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *CommunityVersion*-amzn-*EmrVersion*. O *EmrVersion* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.3.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.3.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.3.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.5.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.16.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.2.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-2	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-2	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-2	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-2	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-2	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.

Componente	Versão	Descrição
hadoop-yarn-nodemanager	2.7.3-amzn-2	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hadoop-yarn-timeline-server	2.7.3-amzn-2	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.3.0	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.3.0	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.3.0	HBase cliente de linha de comando.
hbase-rest-server	1.3.0	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.3.0	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.1.1-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.

Componente	Versão	Descrição
hcatalog-server	2.1.1-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.1.1-amzn-0	Endpoint HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.1.1-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.1.1-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.1.1-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.1.1-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.12.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
oozie-client	4.3.0	O cliente da linha de comando do Oozie.

Componente	Versão	Descrição
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.9.0- -1.2 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.9.0- -1.2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.170	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.170	O serviço que executa partes de uma consulta.
pig-client	0.16.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	2.1.0	Os clientes da linha de comando do Spark.
spark-history-server	2.1.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.1.0	O mecanismo de execução em memória para o YARN.

Componente	Versão	Descrição
spark-yarn-slave	2.1.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.5.4

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.5.4

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.

Classificações	Descrição
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive- site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.pro perties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.pr operties do Hive.

Classificações	Descrição
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.proper ties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.x ml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properti es do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.

Classificações	Descrição
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.

Classificações	Descrição
presto-connector-cassandra	Alterar os valores no arquivo cassandra .properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.p roperties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.propert ies do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.con f do Spark.

Classificações	Descrição
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.5.3 do Amazon EMR

Versões 5.5.3 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGanglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, SparkSqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.5.3	emr-5.5.2	emr-5.5.1	emr-5.5.0
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.2.0	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0

	emr-5.5.3	emr-5.5.2	emr-5.5.1	emr-5.5.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.5.3	emr-5.5.2	emr-5.5.1	emr-5.5.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas da versão 5.5.3

As notas da versão a seguir incluem informações para a versão 5.5.3 do Amazon EMR. As alterações são referentes à versão 5.5.2.

Data da versão inicial: 29 de agosto de 2018

Alterações, melhorias e problemas resolvidos

- Esta versão aborda uma possível vulnerabilidade de segurança.

Versões de componente 5.5.3

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.3.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.3.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.3.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.5.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.16.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.2.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas

Componente	Versão	Descrição
		coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-2	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-2	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-2	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-2	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-2	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-2	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.7.3-amzn-2	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-2	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.3.0	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.3.0	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.3.0	HBase cliente de linha de comando.
hbase-rest-server	1.3.0	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.3.0	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.1.1-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	2.1.1-amzn-0	Endpoint HTTP fornecend o uma interface REST para HCatalog.
hive-client	2.1.1-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.1.1-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.1.1-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.1.1-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.12.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.

Componente	Versão	Descrição
phoenix-library	4.9.0- -1.2 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.9.0- -1.2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.170	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.170	O serviço que executa partes de uma consulta.
pig-client	0.16.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	2.1.0	Os clientes da linha de comando do Spark.
spark-history-server	2.1.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.1.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.1.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.

Componente	Versão	Descrição
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.5.3

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.5.3

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.

Classificações	Descrição
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.

Classificações	Descrição
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.

Classificações	Descrição
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.

Classificações	Descrição
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.

Classificações	Descrição
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.

Classificações	Descrição
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.5.2 do Amazon EMR

Versões 5.5.2 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGanglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, SparkSqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.5.2	emr-5.5.1	emr-5.5.0	emr-5.4.1
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.2.0	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.11.0

	emr-5.5.2	emr-5.5.1	emr-5.5.0	emr-5.4.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.166
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.1	0.7.0
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.9

Notas da versão 5.5.2

As notas da versão a seguir incluem informações para a versão 5.5.2 do Amazon EMR. As alterações são referentes à versão 5.5.1.

Data da versão inicial: 29 de março de 2018

Alterações, melhorias e problemas resolvidos

- Atualizado o kernel do Amazon Linux da AMI padrão do Amazon Linux para Amazon EMR para abordar possíveis vulnerabilidades.

Versões de componente 5.5.2

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.3.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.3.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.3.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.

Componente	Versão	Descrição
emr-s3-dist-cp	2.5.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.16.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.2.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-2	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-2	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-2	O cliente de linha de comando e biblioteca do HDFS

Componente	Versão	Descrição
hadoop-hdfs-namenode	2.7.3-amzn-2	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-2	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-2	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-2	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.3-amzn-2	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-2	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.3.0	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.

Componente	Versão	Descrição
hbase-region-server	1.3.0	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.3.0	HBase cliente de linha de comando.
hbase-rest-server	1.3.0	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.3.0	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.1.1-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.1.1-amzn-0	Endpoint HTTP fornecend o uma interface REST para HCatalog.
hive-client	2.1.1-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.1.1-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.1.1-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.

Componente	Versão	Descrição
hive-server2	2.1.1-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.12.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.9.0- -1.2 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.9.0- -1.2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.170	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.170	O serviço que executa partes de uma consulta.

Componente	Versão	Descrição
pig-client	0.16.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	2.1.0	Os clientes da linha de comando do Spark.
spark-history-server	2.1.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.1.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.1.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.

Componente	Versão	Descrição
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.5.2

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.5.2

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.

Classificações	Descrição
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.

Classificações	Descrição
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.

Classificações	Descrição
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.

Classificações	Descrição
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.

Classificações	Descrição
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.

Classificações	Descrição
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.5.1 do Amazon EMR

Versões 5.5.1 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGanglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, SparkSqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.5.1	emr-5.5.0	emr-5.4.1	emr-5.4.0
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8

	emr-5.5.1	emr-5.5.0	emr-5.4.1	emr-5.4.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.2.0	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.11.0	3.11.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0

	emr-5.5.1	emr-5.5.0	emr-5.4.1	emr-5.4.0
Presto	0.170	0.170	0.166	0.166
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.0	0.7.0
ZooKeeper	3.4.10	3.4.10	3.4.9	3.4.9

Notas da versão 5.5.1

As notas da versão a seguir incluem informações para a versão 5.5.1 do Amazon EMR. As alterações são referentes à versão 5.5.0 do Amazon EMR.

Data da versão inicial: 22 de janeiro de 2018

Alterações, melhorias e problemas resolvidos

- Atualizado o kernel do Amazon Linux da AMI padrão do Amazon Linux para Amazon EMR para abordar vulnerabilidades associadas à execução especulativa (CVE-2017-5715, CVE-2017-5753 e CVE-2017-5754). Para obter mais informações, consulte https://aws.amazon.com/security/security-bulletins/AWS-2018-013/.

Versões de componente 5.5.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a

versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *CommunityVersion*-amzn-*EmrVersion*. O *EmrVersion* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.3.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.3.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.3.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.5.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.16.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.2.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-2	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-2	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-2	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-2	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-2	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.

Componente	Versão	Descrição
hadoop-yarn-nodemanager	2.7.3-amzn-2	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hadoop-yarn-timeline-server	2.7.3-amzn-2	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.3.0	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.3.0	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.3.0	HBase cliente de linha de comando.
hbase-rest-server	1.3.0	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.3.0	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.1.1-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.

Componente	Versão	Descrição
hcatalog-server	2.1.1-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.1.1-amzn-0	Endpoint HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.1.1-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.1.1-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.1.1-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.1.1-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.12.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
oozie-client	4.3.0	O cliente da linha de comando do Oozie.

Componente	Versão	Descrição
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.9.0- -1.2 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.9.0- -1.2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.170	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.170	O serviço que executa partes de uma consulta.
pig-client	0.16.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	2.1.0	Os clientes da linha de comando do Spark.
spark-history-server	2.1.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.1.0	O mecanismo de execução em memória para o YARN.

Componente	Versão	Descrição
spark-yarn-slave	2.1.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.5.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.5.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.

Classificações	Descrição
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive- site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.pro perties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.pr operties do Hive.

Classificações	Descrição
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.proper ties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.x ml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properti es do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.

Classificações	Descrição
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.

Classificações	Descrição
presto-connector-cassandra	Alterar os valores no arquivo cassandra .properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.p roperties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.propert ies do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.con f do Spark.

Classificações	Descrição
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.5.0 do Amazon EMR

Versões 5.5.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGanglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, SparkSqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.5.0	emr-5.4.1	emr-5.4.0	emr-5.3.2
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.2.0	1.2.0	1.1.4
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.2.3

	emr-5.5.0	emr-5.4.1	emr-5.4.0	emr-5.3.2
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.11.0	3.11.0	3.11.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.166	0.166	0.157.1
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.5.0	emr-5.4.1	emr-5.4.0	emr-5.3.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.0	0.7.0	0.6.2
ZooKeeper	3.4.10	3.4.9	3.4.9	3.4.9

Notas da versão 5.5.0

As notas da versão a seguir incluem informações para a versão 5.5.0 do Amazon EMR. As alterações são referentes à versão 5.4.0 do Amazon EMR.

Data do release: 26 de abril de 2017

Atualizações

- Hue 3.12
- Presto 0.170
- Zeppelin 0.7.1
- ZooKeeper 3.4.10

Alterações, melhorias e problemas resolvidos

- Spark
  - Correção do Backported Spark Patch (SPARK-20115) DAGScheduler para recalcular todos os blocos de reprodução aleatória perdidos quando o serviço externo de reprodução aleatória não está disponível na versão 2.1.0 do Spark, incluída nesta versão.
- Flink
  - O Flink agora é compilado com o Scala 2.11. Se você usar a API e as bibliotecas do Scala, recomendamos que use o Scala 2.11 em seus projetos.
  - Tratado um problema em que os padrões `HADOOP_CONF_DIR` e `YARN_CONF_DIR` não estavam definidos corretamente, portanto havia falha no funcionamento de `start-scala-shell.sh`. Também foi adicionada a possibilidade de definir esses valores usando

`env.hadoop.conf.dir` e `env.yarn.conf.dir` em `/etc/flink/conf/flink-conf.yaml` ou na classificação de configuração `flink-conf`.

- Introduzido um novo comando específico do EMR, `flink-scala-shell`, como um wrapper para `start-scala-shell.sh`. Recomendamos o uso desse comando, em vez de `start-scala-shell`. O novo comando simplifica a execução. Por exemplo, `flink-scala-shell -n 2` inicia um shell Scala Flink com um paralelismo de tarefa de 2.
- Introduzido um novo comando específico do EMR, `flink-yarn-session`, como um wrapper para `yarn-session.sh`. Recomendamos o uso desse comando, em vez de `yarn-session`. O novo comando simplifica a execução. Por exemplo, `flink-yarn-session -d -n 2` inicia uma sessão de longa execução do Flink em um estado desanexado com dois gerenciadores de tarefas.
- Tratado (FLINK-6125) Commons httpclient is not shaded anymore in Flink 1.2.

- Presto
  - Adicionado o suporte para autenticação do LDAP. Usar LDAP com o Presto no Amazon EMR requer que você habilite o acesso HTTPS para o coordenador do Presto (`http-server.https.enabled=true` em `config.properties`). Para obter detalhes da configuração, consulte Autenticação LDAP na documentação do Presto.
  - O suporte adicionado para `SHOW GRANTS`.
- AMI base do Linux do Amazon EMR
  - As versões do Amazon EMR agora são baseadas no Amazon Linux 2017.03. Para obter mais informações, consulte Notas da versão da AMI do Amazon Linux 2017.03.
  - Removido o Python 2.6 da imagem de base Linux do Amazon EMR. Python 2.7 e 3.4 estão instalados por padrão. Você pode instalar o Python 2.6 manualmente, se necessário.

Versões de componente 5.5.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`*

começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.3.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.3.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.3.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.5.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.16.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.2.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-2	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-2	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-2	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-2	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-2	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-2	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.7.3-amzn-2	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-2	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.3.0	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.3.0	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.3.0	HBase cliente de linha de comando.
hbase-rest-server	1.3.0	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.3.0	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.1.1-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	2.1.1-amzn-0	Endpoint HTTP fornecend o uma interface REST para HCatalog.
hive-client	2.1.1-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.1.1-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.1.1-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.1.1-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.12.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.

Componente	Versão	Descrição
phoenix-library	4.9.0- -1.2 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.9.0- -1.2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.170	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.170	O serviço que executa partes de uma consulta.
pig-client	0.16.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	2.1.0	Os clientes da linha de comando do Spark.
spark-history-server	2.1.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.1.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.1.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.

Componente	Versão	Descrição
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.10	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.5.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.5.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.

Classificações	Descrição
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.

Classificações	Descrição
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.

Classificações	Descrição
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.

Classificações	Descrição
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.

Classificações	Descrição
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.

Classificações	Descrição
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.4.1 do Amazon EMR

Versões 5.4.1 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGanglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, SparkSqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.4.1	emr-5.4.0	emr-5.3.2	emr-5.3.1
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.2.0	1.1.4	1.1.4
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.2.3	1.2.3
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.11.0	3.11.0	3.11.0	3.11.0

	emr-5.4.1	emr-5.4.0	emr-5.3.2	emr-5.3.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.166	0.166	0.157.1	0.157.1
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.0	0.7.0	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.9

Notas da versão 5.4.1

Esta é uma versão de patch para adicionar a autenticação AWS Signature versão 4 para solicitações ao Amazon S3. Todas as aplicações e componentes são os mesmos que os da versão anterior do Amazon EMR.

> **Important**
>
> Nesta versão de lançamento, o Amazon EMR usa o AWS Signature versão 4 exclusivamente para autenticar solicitações para o Amazon S3. Para obter mais informações, consulte Novidades.

Versões de componente 5.4.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.2.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.3.0	Bibliotecas convenientes para o ecossistema do Hadoop.

Componente	Versão	Descrição
emr-kinesis	3.2.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.15.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.2.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.

Componente	Versão	Descrição
hadoop-hdfs-datanode	2.7.3-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.3-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.

Componente	Versão	Descrição
hbase-hmaster	1.3.0	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.3.0	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.3.0	HBase cliente de linha de comando.
hbase-rest-server	1.3.0	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.3.0	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.1.1-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.1.1-amzn-0	Endpoint HTTP fornecend o uma interface REST para HCatalog.
hive-client	2.1.1-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.1.1-amzn-0	Cliente do Hive-hbase.

Componente	Versão	Descrição
hive-metastore-server	2.1.1-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.1.1-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.11.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.9.0- -1.2 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.9.0- -1.2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica

Componente	Versão	Descrição
presto-coordinator	0.166	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.166	O serviço que executa partes de uma consulta.
pig-client	0.16.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	2.1.0	Os clientes da linha de comando do Spark.
spark-history-server	2.1.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.1.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.1.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.

Componente	Versão	Descrição
zeppelin-server	0.7.0	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.9	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.9	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.4.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.4.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.

Classificações	Descrição
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.

Classificações	Descrição
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.

Classificações	Descrição
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.

Classificações	Descrição
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.

Classificações	Descrição
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.

Classificações	Descrição
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.4.0 do Amazon EMR

Versões 5.4.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGanglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, SparkSqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR

- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.4.0	emr-5.3.2	emr-5.3.1	emr-5.3.0
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.1.4	1.1.4	1.1.4
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.2.3	1.2.3	1.2.3
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.11.0	3.11.0	3.11.0	3.11.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-

	emr-5.4.0	emr-5.3.2	emr-5.3.1	emr-5.3.0
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.166	0.157.1	0.157.1	0.157.1
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.0	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.9

Notas da versão 5.4.0

As notas da versão a seguir incluem informações para a versão 5.4.0 do Amazon EMR. As alterações são referentes à versão 5.3.0 do Amazon EMR.

Data do release: 08 de março de 2017

Atualizações

- Atualizado para Flink 1.2.0

- Atualizado para 1.3.0 HBase
- Atualizado para Phoenix 4.9.0

> **Note**
>
> Se você fez a atualização de uma versão mais antiga do Amazon EMR para a versão 5.4.0 ou posterior do Amazon EMR e usa indexação secundária, atualize os índices locais conforme descrito na documentação do Apache Phoenix. O Amazon EMR remove as configurações necessárias da classificação do `hbase-site`, mas os índices precisam ser preenchidos novamente. O sistema oferece suporte a atualizações de índices online e offline. As atualizações online são o padrão, o que significa que os índices são preenchidos novamente durante a inicialização de clientes do Phoenix versão 4.8.0 ou posterior. Para especificar as atualizações offline, defina a configuração do `phoenix.client.localIndexUpgrade` como falsa na classificação do `phoenix-site` e, em seguida, execute o SSH no nó primário para executar o `psql [zookeeper] -1`.

- Atualizado para Presto 0.166
- Atualizado para Zeppelin 0.7.0

Alterações e melhorias

- Adicionado o suporte para instâncias r4. Consulte os tipos de EC2 instância da Amazon.

Versões de componente 5.4.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado

`myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.2.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.3.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.2.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.15.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.2.0	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.7.3-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.3.0	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.3.0	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.3.0	HBase cliente de linha de comando.
hbase-rest-server	1.3.0	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.3.0	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.1.1-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	2.1.1-amzn-0	Endpoint HTTP fornecend o uma interface REST para HCatalog.
hive-client	2.1.1-amzn-0	O cliente da linha de comando do Hive.
hive-hbase	2.1.1-amzn-0	Cliente do Hive-hbase.
hive-metastore-server	2.1.1-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server2	2.1.1-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.11.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.

Componente	Versão	Descrição
phoenix-library	4.9.0- -1.2 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.9.0- -1.2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.166	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.166	O serviço que executa partes de uma consulta.
pig-client	0.16.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	2.1.0	Os clientes da linha de comando do Spark.
spark-history-server	2.1.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.1.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.1.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.

Componente	Versão	Descrição
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.7.0	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.9	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.9	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.4.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.4.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.

Classificações	Descrição
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.

Classificações	Descrição
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.

Classificações	Descrição
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.

Classificações	Descrição
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.

Classificações	Descrição
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.

Classificações	Descrição
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.3.2 do Amazon EMR

Versões 5.3.2 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGanglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, SparkSqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.3.2	emr-5.3.1	emr-5.3.0	emr-5.2.3
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.4	1.1.4	1.1.4	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3
HCatalog	2.1.1	2.1.1	2.1.1	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.0
Hudi	-	-	-	-
Hue	3.11.0	3.11.0	3.11.0	3.10.0

	emr-5.3.2	emr-5.3.1	emr-5.3.0	emr-5.2.3
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.157.1	0.157.1	0.157.1	0.157.1
Spark	2.1.0	2.1.0	2.1.0	2.0.2
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.9

Notas da versão 5.3.2

Esta é uma versão de patch para adicionar a autenticação AWS Signature versão 4 para solicitações ao Amazon S3. Todas as aplicações e componentes são os mesmos que os da versão anterior do Amazon EMR.

> **Important**
>
> Nesta versão de lançamento, o Amazon EMR usa o AWS Signature versão 4 exclusivamente para autenticar solicitações para o Amazon S3. Para obter mais informações, consulte Novidades.

Versões de componente 5.3.2

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.2.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.2.0	Bibliotecas convenientes para o ecossistema do Hadoop.

Componente	Versão	Descrição
emr-kinesis	3.2.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.14.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.1.4	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.

Componente	Versão	Descrição
hadoop-hdfs-datanode	2.7.3-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.3-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.

Componente	Versão	Descrição
hbase-hmaster	1.2.3	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.2.3	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.3	HBase cliente de linha de comando.
hbase-rest-server	1.2.3	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.3	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.1.1-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.1.1-amzn-0	Endpoint HTTP fornecend o uma interface REST para HCatalog.
hive-client	2.1.1-amzn-0	O cliente da linha de comando do Hive.

Componente	Versão	Descrição
hive-metastore-server	2.1.1-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	2.1.1-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.11.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.7.0- -1,2 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.7.0- -1,2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica

Componente	Versão	Descrição
presto-coordinator	0.157.1	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.157.1	O serviço que executa partes de uma consulta.
pig-client	0.16.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	2.1.0	Os clientes da linha de comando do Spark.
spark-history-server	2.1.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.1.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.1.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.

Componente	Versão	Descrição
zeppelin-server	0.6.2	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.9	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.9	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.3.2

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.3.2

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.

Classificações	Descrição
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.

Classificações	Descrição
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.

Classificações	Descrição
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.

Classificações	Descrição
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.

Classificações	Descrição
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.

Classificações	Descrição
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.3.1 do Amazon EMR

Versões 5.3.1 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGanglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, SparkSqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR

- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.3.1	emr-5.3.0	emr-5.2.3	emr-5.2.2
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.4	1.1.4	1.1.3	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3
HCatalog	2.1.1	2.1.1	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.11.0	3.11.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-

	emr-5.3.1	emr-5.3.0	emr-5.2.3	emr-5.2.2
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.157.1	0.157.1	0.157.1	0.157.1
Spark	2.1.0	2.1.0	2.0.2	2.0.2
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.9

Notas da versão 5.3.1

As notas da versão a seguir incluem informações para a versão 5.3.1 do Amazon EMR. As alterações são referentes à versão 5.3.0 do Amazon EMR.

Data da versão: 7 de fevereiro de 2017

Pequenas alterações nos patches do Zeppelin de backport e na atualização da AMI padrão para o Amazon EMR.

Versões de componente 5.3.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.2.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.2.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.2.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.14.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.1.4	Scripts do cliente da linha de comando e aplicativos do Apache Flink.

Componente	Versão	Descrição
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-1	O endpoint de HTTP para as operações do HDFS.

Componente	Versão	Descrição
hadoop-kms-server	2.7.3-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.3-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.2.3	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.2.3	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.3	HBase cliente de linha de comando.

Componente	Versão	Descrição
hbase-rest-server	1.2.3	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.3	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.1.1-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.1.1-amzn-0	Endpoint HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.1.1-amzn-0	O cliente da linha de comando do Hive.
hive-metastore-server	2.1.1-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	2.1.1-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.11.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop

Componente	Versão	Descrição
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.7.0- -1,2 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.7.0- -1,2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.157.1	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.157.1	O serviço que executa partes de uma consulta.
pig-client	0.16.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	2.1.0	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	2.1.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.1.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.1.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.6.2	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.9	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.9	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.3.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.3.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop

Classificações	Descrição
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive- site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.

Classificações	Descrição
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.

Classificações	Descrição
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.

Classificações	Descrição
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.

Classificações	Descrição
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.

Classificações	Descrição
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.3.0 do Amazon EMR

Versões 5.3.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGanglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, SparkSqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.3.0	emr-5.2.3	emr-5.2.2	emr-5.2.1
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-5.3.0	emr-5.2.3	emr-5.2.2	emr-5.2.1
Flink	1.1.4	1.1.3	1.1.3	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3
HCatalog	2.1.1	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.11.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.157.1	0.157.1	0.157.1	0.157.1
Spark	2.1.0	2.0.2	2.0.2	2.0.2
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6

	emr-5.3.0	emr-5.2.3	emr-5.2.2	emr-5.2.1
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.9

Notas da versão 5.3.0

As notas da versão a seguir incluem informações para a versão 5.3.0 do Amazon EMR. As alterações são referentes à versão 5.2.1 do Amazon EMR.

Data do release: 26 de janeiro de 2017

Atualizações

- Atualizado para Hive 2.1.1
- Atualizado para Hue 3.11.0
- Atualizado para Spark 2.1.0
- Atualizado para Oozie 4.3.0
- Atualizado para Flink 1.1.4

Alterações e melhorias

- Adicionado um patch para o Hue que permite usar a configuração `interpreters_shown_on_wheel` para definir o que intérpretes mostram primeiro na roda de seleção do bloco de anotações, independentemente de sua ordem no arquivo `hue.ini`.
- Adicionada a classificação de configuração `hive-parquet-logging`, que pode ser usada para configurar os valores no arquivo `parquet-logging.properties` do Hive.

Versões de componente 5.3.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.2.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.2.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.2.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.14.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.1.4	Scripts do cliente da linha de comando e aplicativos do Apache Flink.

Componente	Versão	Descrição
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-1	O endpoint de HTTP para as operações do HDFS.

Componente	Versão	Descrição
hadoop-kms-server	2.7.3-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.3-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.2.3	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.2.3	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.3	HBase cliente de linha de comando.

Componente	Versão	Descrição
hbase-rest-server	1.2.3	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.3	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.1.1-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.1.1-amzn-0	Endpoint HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.1.1-amzn-0	O cliente da linha de comando do Hive.
hive-metastore-server	2.1.1-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	2.1.1-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.11.0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop

Componente	Versão	Descrição
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.52	O servidor de banco de dados MySQL.
oozie-client	4.3.0	O cliente da linha de comando do Oozie.
oozie-server	4.3.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.7.0- -1,2 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.7.0- -1,2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.157.1	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.157.1	O serviço que executa partes de uma consulta.
pig-client	0.16.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	2.1.0	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	2.1.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.1.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.1.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.6.2	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.9	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.9	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.3.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.3.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop

Classificações	Descrição
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive- site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.

Classificações	Descrição
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-parquet-logging	Alterar os valores no arquivo parquet-logging.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.

Classificações	Descrição
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.

Classificações	Descrição
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.

Classificações	Descrição
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.

Classificações	Descrição
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.2.3 do Amazon EMR

Versões 5.2.3 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGanglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, SparkSqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.2.3	emr-5.2.2	emr-5.2.1	emr-5.2.0
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-5.2.3	emr-5.2.2	emr-5.2.1	emr-5.2.0
Flink	1.1.3	1.1.3	1.1.3	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.157.1	0.157.1	0.157.1	0.152.3
Spark	2.0.2	2.0.2	2.0.2	2.0.2
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6

	emr-5.2.3	emr-5.2.2	emr-5.2.1	emr-5.2.0
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.8

Notas da versão 5.2.3

Esta é uma versão de patch para adicionar a autenticação AWS Signature versão 4 para solicitações ao Amazon S3. Todas as aplicações e componentes são os mesmos que os da versão anterior do Amazon EMR.

> **Important**
>
> Nesta versão de lançamento, o Amazon EMR usa o AWS Signature versão 4 exclusivamente para autenticar solicitações para o Amazon S3. Para obter mais informações, consulte Novidades.

Versões de componente 5.2.3

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado

`myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.2.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.2.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.2.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.13.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.1.3	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.7.3-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.2.3	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.2.3	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.3	HBase cliente de linha de comando.
hbase-rest-server	1.2.3	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.3	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.1.0-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.1.0-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	2.1.0-amzn-0	Endpoint HTTP fornecend o uma interface REST para HCatalog.
hive-client	2.1.0-amzn-0	O cliente da linha de comando do Hive.
hive-metastore-server	2.1.0-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	2.1.0-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.10.0-amzn-0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.52	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.7.0- -1,2 HBase	As bibliotecas do phoenix para servidor e cliente

Componente	Versão	Descrição
phoenix-query-server	4.7.0- -1,2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.157.1	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.157.1	O serviço que executa partes de uma consulta.
pig-client	0.16.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	2.0.2	Os clientes da linha de comando do Spark.
spark-history-server	2.0.2	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.0.2	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.0.2	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.

Componente	Versão	Descrição
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.23	O servidor Apache HTTP.
zeppelin-server	0.6.2	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.9	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.9	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.2.3

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.2.3

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.

Classificações	Descrição
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.

Classificações	Descrição
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.

Classificações	Descrição
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.

Classificações	Descrição
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.

Classificações	Descrição
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.

Classificações	Descrição
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.2.2 do Amazon EMR

Versões 5.2.2 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGanglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, SparkSqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR

- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.2.2	emr-5.2.1	emr-5.2.0	emr-5.1.1
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.3	1.1.3	1.1.3	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-5.2.2	emr-5.2.1	emr-5.2.0	emr-5.1.1
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.157.1	0.157.1	0.152.3	0.152.3
Spark	2.0.2	2.0.2	2.0.2	2.0.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.8	3.4.8

Notas da versão 5.2.2

As notas da versão a seguir incluem informações para a versão 5.2.2 do Amazon EMR. As alterações são referentes à versão 5.2.1 do Amazon EMR.

Data do release: 2 de maio de 2017

Problemas conhecidos das versões anteriores que foram resolvidos

- Backported SPARK-194459, que soluciona um problema em que a leitura de uma tabela ORC com colunas pode falhar. char/varchar

Versões de componente 5.2.2

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.2.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.2.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.2.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.

Componente	Versão	Descrição
emrfs	2.13.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
flink-client	1.1.3	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-1	O cliente de linha de comando e biblioteca do HDFS

Componente	Versão	Descrição
hadoop-hdfs-namenode	2.7.3-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.3-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.2.3	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.

Componente	Versão	Descrição
hbase-region-server	1.2.3	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.3	HBase cliente de linha de comando.
hbase-rest-server	1.2.3	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.3	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.1.0-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.1.0-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.1.0-amzn-0	Endpoint HTTP fornecend o uma interface REST para HCatalog.
hive-client	2.1.0-amzn-0	O cliente da linha de comando do Hive.
hive-metastore-server	2.1.0-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.

Componente	Versão	Descrição
hive-server	2.1.0-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.10.0-amzn-0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.52	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.7.0- -1,2 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.7.0- -1,2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.157.1	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.157.1	O serviço que executa partes de uma consulta.

Componente	Versão	Descrição
pig-client	0.16.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	2.0.2	Os clientes da linha de comando do Spark.
spark-history-server	2.0.2	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.0.2	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.0.2	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.23	O servidor Apache HTTP.
zeppelin-server	0.6.2	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.9	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.

Componente	Versão	Descrição
zookeeper-client	3.4.9	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.2.2

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.2.2

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.

Classificações	Descrição
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.

Classificações	Descrição
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.

Classificações	Descrição
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.

Classificações	Descrição
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.

Classificações	Descrição
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.

Classificações	Descrição
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.2.1 do Amazon EMR

Versões 5.2.1 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGanglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, SparkSqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.2.1	emr-5.2.0	emr-5.1.1	emr-5.1.0
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8

	emr-5.2.1	emr-5.2.0	emr-5.1.1	emr-5.1.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.3	1.1.3	1.1.3	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0

	emr-5.2.1	emr-5.2.0	emr-5.1.1	emr-5.1.0
Presto	0.157.1	0.152.3	0.152.3	0.152.3
Spark	2.0.2	2.0.2	2.0.1	2.0.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.8	3.4.8	3.4.8

Notas da versão 5.2.1

As notas da versão a seguir incluem informações para a versão 5.2.1 do Amazon EMR. As alterações são referentes à versão 5.2.0 do Amazon EMR.

Data do release: 29 de dezembro de 2016

Atualizações

- Atualizado para Presto 0.157.1. Para obter mais informações, consulte Notas de release do Presto na documentação do Presto.
- Zookeeper atualizado para 3.4.9. Para obter mais informações, consulte as notas de ZooKeeper lançamento na ZooKeeper documentação do Apache.

Alterações e melhorias

- Foi adicionado suporte para o tipo de instância Amazon EC2 m4.16xlarge no Amazon EMR versão 4.8.3 e posterior, excluindo 5.0.0, 5.0.3 e 5.2.0.
- As versões do Amazon EMR agora são baseadas no Amazon Linux 2016.09. Para obter mais informações, consulte https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/.

- A localização de caminhos de configuração do Flink e do YARN agora são definidas por padrão em `/etc/default/flink` para que você não precise definir as variáveis de ambiente `FLINK_CONF_DIR` e `HADOOP_CONF_DIR` ao executar os scripts de driver `flink` ou `yarn-session.sh` para iniciar trabalhos do Flink.
- Foi adicionado suporte para a FlinkKinesisConsumer aula.

Problemas conhecidos das versões anteriores que foram resolvidos

- Corrigido um problema no Hadoop em que o ReplicationMonitor thread podia ficar preso por um longo tempo devido a uma corrida entre a replicação e a exclusão do mesmo arquivo em um grande cluster.
- Corrigido um problema em que ControlledJob #toString falhava com uma exceção de ponteiro nulo (NPE) quando o status do trabalho não era atualizado com êxito.

Versões de componente 5.2.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.2.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.2.0	Bibliotecas convenientes para o ecossistema do Hadoop.

Componente	Versão	Descrição
emr-kinesis	3.2.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.13.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.1.3	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.

Componente	Versão	Descrição
hadoop-hdfs-datanode	2.7.3-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.3-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.

Componente	Versão	Descrição
hbase-hmaster	1.2.3	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.2.3	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.3	HBase cliente de linha de comando.
hbase-rest-server	1.2.3	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.3	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.1.0-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.1.0-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.1.0-amzn-0	Endpoint HTTP fornecend o uma interface REST para HCatalog.
hive-client	2.1.0-amzn-0	O cliente da linha de comando do Hive.

Componente	Versão	Descrição
hive-metastore-server	2.1.0-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	2.1.0-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.10.0-amzn-0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.52	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.7.0- -1,2 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.7.0- -1,2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica

Componente	Versão	Descrição
presto-coordinator	0.157.1	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.157.1	O serviço que executa partes de uma consulta.
pig-client	0.16.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	2.0.2	Os clientes da linha de comando do Spark.
spark-history-server	2.0.2	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.0.2	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.0.2	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.23	O servidor Apache HTTP.

Componente	Versão	Descrição
zeppelin-server	0.6.2	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.9	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.9	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.2.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.2.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.

Classificações	Descrição
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.

Classificações	Descrição
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.

Classificações	Descrição
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.

Classificações	Descrição
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.

Classificações	Descrição
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.

Classificações	Descrição
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.2.0 do Amazon EMR

Versões 5.2.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGanglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, SparkSqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.2.0	emr-5.1.1	emr-5.1.0	emr-5.0.3
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.3	1.1.3	1.1.3	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-5.2.0	emr-5.1.1	emr-5.1.0	emr-5.0.3
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.152.3	0.152.3	0.152.3	0.152.3
Spark	2.0.2	2.0.1	2.0.1	2.0.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

Notas da versão 5.2.0

As notas da versão a seguir incluem informações para a versão 5.2.0 do Amazon EMR. As alterações são referentes à versão 5.1.0 do Amazon EMR.

Data do release: 21 de novembro de 2016

Alterações e melhorias

- Adicionado o modo de armazenamento Amazon S3 para. HBase
- Permite que você especifique uma localização do Amazon S3 para o HBase rootdir. Para obter mais informações, consulte HBase no Amazon S3.

Atualizações

- Atualizado para Spark 2.0.2

Problemas conhecidos das versões anteriores que foram resolvidos

- Corrigido um problema com /mnt ficar restrito a 2 TB em tipos de instância somente EBS.
- Corrigido um problema com o controlador de instância e com os logs do logpusher serem a saída para seus arquivos .out correspondentes, em vez de para seus arquivos normais .log configurados com log4j, que mudam de hora em hora. Os arquivos .out não alternam, portanto isso acabará por encher a partição /emr. Esse problema afeta somente tipos de instância de HVM (máquina virtual de hardware).

Versões de componente 5.2.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.1.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.1.0	Bibliotecas convenientes para o ecossistema do Hadoop.

Componente	Versão	Descrição
emr-kinesis	3.2.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.12.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.1.3	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-0	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.

Componente	Versão	Descrição
hadoop-hdfs-datanode	2.7.3-amzn-0	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-0	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-0	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-0	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-0	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-0	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-0	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.3-amzn-0	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-0	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.

Componente	Versão	Descrição
hbase-hmaster	1.2.3	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.2.3	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.3	HBase cliente de linha de comando.
hbase-rest-server	1.2.3	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.3	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.1.0-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.1.0-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.1.0-amzn-0	Endpoint HTTP fornecend o uma interface REST para HCatalog.
hive-client	2.1.0-amzn-0	O cliente da linha de comando do Hive.

Componente	Versão	Descrição
hive-metastore-server	2.1.0-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	2.1.0-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.10.0-amzn-0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.52	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.7.0- -1,2 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.7.0- -1,2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica

Componente	Versão	Descrição
presto-coordinator	0.152.3	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.152.3	O serviço que executa partes de uma consulta.
pig-client	0.16.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	2.0.2	Os clientes da linha de comando do Spark.
spark-history-server	2.0.2	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.0.2	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.0.2	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.23	O servidor Apache HTTP.

Componente	Versão	Descrição
zeppelin-server	0.6.2	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.8	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.2.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.2.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.

Classificações	Descrição
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase	Configurações selecionadas pelo Amazon EMR para Apache. HBase
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.

Classificações	Descrição
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.

Classificações	Descrição
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.

Classificações	Descrição
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.

Classificações	Descrição
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.

Classificações	Descrição
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.1.1 do Amazon EMR

Versões 5.1.1 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGanglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, SparkSqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.1.1	emr-5.1.0	emr-5.0.3	emr-5.0.2
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.3	1.1.3	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.2	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-5.1.1	emr-5.1.0	emr-5.0.3	emr-5.0.2
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.152.3	0.152.3	0.152.3	0.150
Spark	2.0.1	2.0.1	2.0.1	2.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

Notas da versão 5.1.1

Esta é uma versão de patch para adicionar a autenticação AWS Signature versão 4 para solicitações ao Amazon S3. Todas as aplicações e componentes são os mesmos que os da versão anterior do Amazon EMR.

> **Important**
>
> Nesta versão de lançamento, o Amazon EMR usa o AWS Signature versão 4 exclusivamente para autenticar solicitações para o Amazon S3. Para obter mais informações, consulte Novidades.

Versões de componente 5.1.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.1.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.1.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.2.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.11.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.1.3	Scripts do cliente da linha de comando e aplicativos do Apache Flink.

Componente	Versão	Descrição
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-0	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-0	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-0	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-0	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-0	O endpoint de HTTP para as operações do HDFS.

Componente	Versão	Descrição
hadoop-kms-server	2.7.3-amzn-0	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-0	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-0	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.3-amzn-0	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-0	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.2.3	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.2.3	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.3	HBase cliente de linha de comando.

Componente	Versão	Descrição
hbase-rest-server	1.2.3	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.3	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.1.0-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.1.0-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.1.0-amzn-0	Endpoint HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.1.0-amzn-0	O cliente da linha de comando do Hive.
hive-metastore-server	2.1.0-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	2.1.0-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.10.0-amzn-0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop

Componente	Versão	Descrição
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.52	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.7.0- -1,2 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.7.0- -1,2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.152.3	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.152.3	O serviço que executa partes de uma consulta.
pig-client	0.16.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	2.0.1	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	2.0.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.0.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.0.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.23	O servidor Apache HTTP.
zeppelin-server	0.6.2	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.8	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.1.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.1.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-session .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop

Classificações	Descrição
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.pro perties do Hive.

Classificações	Descrição
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log 4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.proper ties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.x ml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properti es do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.

Classificações	Descrição
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.

Classificações	Descrição
presto-connector-cassandra	Alterar os valores no arquivo cassandra .properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.p roperties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.propert ies do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.con f do Spark.

Classificações	Descrição
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.1.0 do Amazon EMR

Versões 5.1.0 da aplicação

Essa versão inclui os seguintes aplicativos: FlinkGanglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, SparkSqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.1.0	emr-5.0.3	emr-5.0.2	emr-5.0.1
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.3	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.2	1.2.2	1.2.2

	emr-5.1.0	emr-5.0.3	emr-5.0.2	emr-5.0.1
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.2	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.152.3	0.152.3	0.150	0.150
Spark	2.0.1	2.0.1	2.0.0	2.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.1.0	emr-5.0.3	emr-5.0.2	emr-5.0.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.1	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

Notas da versão 5.1.0

As notas da versão a seguir incluem informações para a versão 5.1.0 do Amazon EMR. As alterações são referentes à versão 5.0.3 do Amazon EMR.

Data do release: 03 de novembro de 2016

Alterações e melhorias

- Adicionado o suporte para Flink 1.1.3.
- O Presto foi adicionado como uma opção na seção bloco de anotações do Hue.

Atualizações

- Atualizado para 1.2.3 HBase
- Atualizado para Zeppelin 0.6.2

Problemas conhecidos das versões anteriores que foram resolvidos

- Corrigido um problema com consultas do Tez no Amazon S3 com arquivos ORC que não tinham uma performance tão boa quanto em versões 4.x anteriores do Amazon EMR.

Versões de componente 5.1.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a

versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *CommunityVersion*-amzn-*EmrVersion*. O *EmrVersion* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.1.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.1.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.2.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.11.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
flink-client	1.1.3	Scripts do cliente da linha de comando e aplicativos do Apache Flink.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-0	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-0	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-0	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-0	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-0	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-0	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.7.3-amzn-0	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.

Componente	Versão	Descrição
hadoop-yarn-nodemanager	2.7.3-amzn-0	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-0	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hadoop-yarn-timeline-server	2.7.3-amzn-0	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.2.3	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.2.3	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.3	HBase cliente de linha de comando.
hbase-rest-server	1.2.3	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.3	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.1.0-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.

Componente	Versão	Descrição
hcatalog-server	2.1.0-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.1.0-amzn-0	Endpoint HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.1.0-amzn-0	O cliente da linha de comando do Hive.
hive-metastore-server	2.1.0-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	2.1.0-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.10.0-amzn-0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.52	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.

Componente	Versão	Descrição
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.7.0- -1,2 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.7.0- -1,2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.152.3	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.152.3	O serviço que executa partes de uma consulta.
pig-client	0.16.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	2.0.1	Os clientes da linha de comando do Spark.
spark-history-server	2.0.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.0.1	O mecanismo de execução em memória para o YARN.

Componente	Versão	Descrição
spark-yarn-slave	2.0.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.23	O servidor Apache HTTP.
zeppelin-server	0.6.2	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.8	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.1.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.1.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
flink-conf	Alterar as configurações do flink-conf.yaml.
flink-log4j	Alterar as configurações de log4j.properties no Flink.
flink-log4j-yarn-session	Altere as configurações do Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Alterar as configurações de log4j-cli.properties no Flink.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.

Classificações	Descrição
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.

Classificações	Descrição
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.proper ties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.x ml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properti es do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.

Classificações	Descrição
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.

Classificações	Descrição
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.

Classificações	Descrição
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.0.3 do Amazon EMR

Versões 5.0.3 da aplicação

Esta versão inclui os seguintes aplicativos: GangliaHBase, HCatalog, Hadoop, Hive, Hue, Mahout, Oozie, Phoenix, PigPresto, Spark, Sqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.2	2.7.2	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.152.3	0.150	0.150	0.150
Spark	2.0.1	2.0.0	2.0.0	2.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

Notas da versão 5.0.3

As notas da versão a seguir incluem informações para a versão 5.0.3 do Amazon EMR. As alterações são referentes à versão 5.0.0 do Amazon EMR.

Data do release: 24 de outubro de 2016

Atualizações

- Atualizado para Hadoop 2.7.3
- Atualizado para Presto 0.152.3, que inclui o suporte para a interface da web do Presto. Você pode acessar a interface da web do Presto no coordenador do Presto usando a porta 8889. Para obter mais informações sobre a interface da web do Presto, consulte Interface da web na documentação do Presto.
- Atualizado para Spark 2.0.1
- As versões do Amazon EMR agora são baseadas no Amazon Linux 2016.09. Para obter mais informações, consulte https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/.

Versões de componente 5.0.3

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.1.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.

Componente	Versão	Descrição
emr-goodies	2.1.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.2.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.10.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-0	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-0	O serviço de nível de nó do HDFS para armazenamento de blocos.

Componente	Versão	Descrição
hadoop-hdfs-library	2.7.3-amzn-0	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-0	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-0	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-0	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-0	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-0	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.3-amzn-0	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-0	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.

Componente	Versão	Descrição
hbase-hmaster	1.2.2	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.2.2	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.2	HBase cliente de linha de comando.
hbase-rest-server	1.2.2	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.2	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.1.0-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.1.0-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	2.1.0-amzn-0	Endpoint HTTP fornecend o uma interface REST para HCatalog.
hive-client	2.1.0-amzn-0	O cliente da linha de comando do Hive.

Componente	Versão	Descrição
hive-metastore-server	2.1.0-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	2.1.0-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.10.0-amzn-0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.52	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.7.0- -1,2 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.7.0- -1,2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica

Componente	Versão	Descrição
presto-coordinator	0.152.3	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.152.3	O serviço que executa partes de uma consulta.
pig-client	0.16.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	2.0.1	Os clientes da linha de comando do Spark.
spark-history-server	2.0.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.0.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.0.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.23	O servidor Apache HTTP.

Componente	Versão	Descrição
zeppelin-server	0.6.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.8	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.0.3

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.0.3

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.

Classificações	Descrição
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive- site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.

Classificações	Descrição
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.

Classificações	Descrição
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.

Classificações	Descrição
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.

Classificações	Descrição
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.

Classificações	Descrição
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.0.2 do Amazon EMR

Versões 5.0.2 da aplicação

Esta versão inclui os seguintes aplicativos: GangliaHBase, HCatalog, Hadoop, Hive, Hue, Mahout, Oozie, Phoenix, PigPresto, Spark, Sqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.2	2.7.2	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.152.3	0.150	0.150	0.150
Spark	2.0.1	2.0.0	2.0.0	2.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

Notas da versão 5.0.2

Versões de componente 5.0.2

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.0.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.1.0	Bibliotecas convenientes para o ecossistema do Hadoop.

Componente	Versão	Descrição
emr-kinesis	3.2.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.9.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.2-amzn-3	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.2-amzn-3	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.2-amzn-3	O cliente de linha de comando e biblioteca do HDFS

Componente	Versão	Descrição
hadoop-hdfs-namenode	2.7.2-amzn-3	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.2-amzn-3	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.2-amzn-3	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.2-amzn-3	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.2-amzn-3	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.2-amzn-3	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.2-amzn-3	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.2.2	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.

Componente	Versão	Descrição
hbase-region-server	1.2.2	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.2	HBase cliente de linha de comando.
hbase-rest-server	1.2.2	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.2	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.1.0-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.1.0-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.1.0-amzn-0	Endpoint HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.1.0-amzn-0	O cliente da linha de comando do Hive.
hive-metastore-server	2.1.0-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.

Componente	Versão	Descrição
hive-server	2.1.0-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.10.0-amzn-0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.46	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.7.0- -1,2 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.7.0- -1,2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.150	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.150	O serviço que executa partes de uma consulta.

Componente	Versão	Descrição
pig-client	0.16.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	2.0.0	Os clientes da linha de comando do Spark.
spark-history-server	2.0.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.0.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.0.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.23	O servidor Apache HTTP.
zeppelin-server	0.6.1-SNAPSHOT	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.8	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.

Componente	Versão	Descrição
zookeeper-client	3.4.8	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.0.2

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.0.2

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.

Classificações	Descrição
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.

Classificações	Descrição
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log 4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.proper ties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.x ml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properti es do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.

Classificações	Descrição
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.

Classificações	Descrição
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.

Classificações	Descrição
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.0.1 do Amazon EMR

Versões 5.0.1 da aplicação

Esta versão inclui os seguintes aplicativos: GangliaHBase, HCatalog, Hadoop, Hive, Hue, Mahout, Oozie, Phoenix, PigPresto, Spark, Sqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.2	2.7.2	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.152.3	0.150	0.150	0.150
Spark	2.0.1	2.0.0	2.0.0	2.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

Notas da versão 5.0.1

Versões de componente 5.0.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.0.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.1.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.2.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.9.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do

Componente	Versão	Descrição
		ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.2-amzn-3	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.2-amzn-3	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.2-amzn-3	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.2-amzn-3	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.2-amzn-3	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.2-amzn-3	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er

Componente	Versão	Descrição
hadoop-mapred	2.7.2-amzn-3	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.2-amzn-3	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.7.2-amzn-3	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hadoop-yarn-timeline-server	2.7.2-amzn-3	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.2.2	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.2.2	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.2	HBase cliente de linha de comando.
hbase-rest-server	1.2.2	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.2	Serviço que fornece um endpoint Thrift para. HBase

Componente	Versão	Descrição
hcatalog-client	2.1.0-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.1.0-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.1.0-amzn-0	Endpoint HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.1.0-amzn-0	O cliente da linha de comando do Hive.
hive-metastore-server	2.1.0-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	2.1.0-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.10.0-amzn-0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.46	O servidor de banco de dados MySQL.

Componente	Versão	Descrição
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.7.0- -1,2 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.7.0- -1,2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.150	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.150	O serviço que executa partes de uma consulta.
pig-client	0.16.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	2.0.0	Os clientes da linha de comando do Spark.
spark-history-server	2.0.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.

Componente	Versão	Descrição
spark-on-yarn	2.0.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.0.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.23	O servidor Apache HTTP.
zeppelin-server	0.6.1-SNAPSHOT	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.8	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.0.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.0.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altere os valores no HBase arquivo hbase-sit e.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.

Classificações	Descrição
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.

Classificações	Descrição
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.

Classificações	Descrição
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.

Classificações	Descrição
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.

Classificações	Descrição
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versão 5.0.0 do Amazon EMR

Versões 5.0.0 da aplicação

Esta versão inclui os seguintes aplicativos: GangliaHBase, HCatalog, Hadoop, Hive, Hue, Mahout, Oozie, Phoenix, PigPresto, Spark, Sqoop, Tez, Zeppelin,, ZooKeepere.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.2	2.7.2	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.152.3	0.150	0.150	0.150
Spark	2.0.1	2.0.0	2.0.0	2.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

Notas da versão 5.0.0

Data do release: 27 de julho de 2016

Atualizações

- Atualizado para Hive 2.1
- Atualizado para Presto 0.150
- Atualizado para Spark 2.0
- Atualizado para Hue 3.10.0
- Atualizado para Pig 0.16.0
- Atualizado para Tez 0.8.4

- Atualizado para Zeppelin 0.6.1

Alterações e melhorias

- O Amazon EMR é compatível com as versões de código aberto mais recentes do Hive (versão 2.1) e do Pig (versão 0.16.0). Se, anteriormente, você tiver usado o Hive ou Pig no Amazon EMR, isso poderá afetar alguns casos de uso. Para obter mais informações, consulte Hive e Pig.
- O mecanismo de execução padrão para o Hive e o Pig agora é o Tez. Para alterar isso, você deve editar os valores apropriados nas classificações de configuração `hive-site` e `pig-properties`, respectivamente.
- Um recurso de etapa aprimorada de depuração foi adicionado, o que permite que você veja a causa raiz de falhas de etapa se o serviço puder determinar a causa. Para obter mais informações, consulte Depuração de etapa aprimorada no Guia de gerenciamento do Amazon EMR.
- Os aplicativos que, anteriormente, terminavam com "-Sandbox" não têm mais esse sufixo. Isso pode inutilizar sua automação, por exemplo, se você estiver usando scripts para iniciar clusters com esses aplicativos. A tabela a seguir mostra nomes de aplicações na versão 4.7.2 do Amazon EMR em relação à versão 5.0.0 do Amazon EMR.

  Alterações dos nomes de aplicativos

Amazon EMR 4.7.2	Amazon EMR 5.0.0
Oozie-Sandbox	Oozie
Presto-Sandbox	Presto
Sqoop-Sandbox	Sqoop
Zeppelin-Sandbox	Zeppelin
ZooKeeper-Caixa de areia	ZooKeeper

- O Spark agora está compilado para Scala 2.11.
- Java 8 é agora o JVM padrão. Todas as aplicações são executadas usando o runtime do Java 8. Não há alterações em qualquer destino de código de bytes da aplicação. A maioria dos aplicativos continuam a usar o Java 7 como destino.
- O Zeppelin agora inclui recursos de autenticação. Para obter mais informações, consulte Zeppelin.

- Adicionado o suporte para configurações de segurança, que permitem criar e aplicar opções de criptografia com mais facilidade. Para obter mais informações, consulte Criptografia de dados.

Versões de componente 5.0.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.0.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.1.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.2.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.9.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.

Componente	Versão	Descrição
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.2-amzn-3	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.2-amzn-3	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.2-amzn-3	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.2-amzn-3	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.2-amzn-3	O endpoint de HTTP para as operações do HDFS.

Componente	Versão	Descrição
hadoop-kms-server	2.7.2-amzn-3	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.2-amzn-3	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.2-amzn-3	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.2-amzn-3	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.2-amzn-3	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.2.2	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.2.2	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.2	HBase cliente de linha de comando.

Componente	Versão	Descrição
hbase-rest-server	1.2.2	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.2	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	2.1.0-amzn-0	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	2.1.0-amzn-0	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	2.1.0-amzn-0	Endpoint HTTP fornecendo uma interface REST para HCatalog.
hive-client	2.1.0-amzn-0	O cliente da linha de comando do Hive.
hive-metastore-server	2.1.0-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	2.1.0-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.10.0-amzn-0	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop

Componente	Versão	Descrição
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.46	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.7.0- -1,2 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.7.0- -1,2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.150	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.150	O serviço que executa partes de uma consulta.
pig-client	0.16.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	2.0.0	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	2.0.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	2.0.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	2.0.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.23	O servidor Apache HTTP.
zeppelin-server	0.6.1-SNAPSHOT	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.8	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de linha de comando.

Classificações de configuração 5.0.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-5.0.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.

Classificações	Descrição
hbase-site	Altere os valores no HBase arquivo hbase-site.xml do.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Altere valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j2	Altere os valores nas propriedades HCat log4j2.properties da HCatalog Web.
hcatalog-webhcat-site	Altere os valores no arquivo webhcat-site.xml HCat da HCatalog Web.
hive-beeline-log4j2	Alterar os valores no arquivo beeline-log4j2.properties do Hive.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j2	Altere os valores no arquivo hive-exec-log4j2.properties do Hive.
hive-llap-daemon-log4j2	Altere os valores no arquivo llap-daemon-log4j2.properties do Hive.
hive-log4j2	Alterar os valores no arquivo hive-log4j2.properties do Hive.

Classificações	Descrição
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altere os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.

Classificações	Descrição
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.

Classificações	Descrição
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-hive-site	Alterar os valores no arquivo hive-site.xml do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.

Classificações	Descrição
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altere os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altere os valores no ZooKeeper arquivo log4j.properties.

Versões de lançamento 4.x do Amazon EMR

Esta seção contém versões de aplicações, notas da versão, versões de componentes e classificações de configuração disponíveis em cada versão de lançamento 4.x do Amazon EMR.

Ao executar um cluster, você pode escolher entre várias versões do Amazon EMR. Isso permite que você teste e use versões de aplicações que atendem aos seus requisitos de compatibilidade. Você especifica o número da versão com o rótulo da versão. Os rótulos de versão estão no formato `emr-x.x.x`. Por exemplo, `.emr-7.9.0`

Novas versões do Amazon EMR são disponibilizadas em diferentes regiões durante um período de vários anos, começando com a primeira região na data da versão inicial. A versão mais recente pode não estar disponível em sua região durante esse período.

Para obter uma tabela abrangente das versões de aplicações em cada versão 4.x do Amazon EMR, consulte Versões de aplicações nas versões 4.x do Amazon EMR.

Tópicos

- Versões de aplicações nas versões 4.x do Amazon EMR
- Diferenças nas versões 4.x do Amazon EMR
- Versão 4.9.6 do Amazon EMR
- Versão 4.9.5 do Amazon EMR
- Versão 4.9.4 do Amazon EMR
- Versão 4.9.3 do Amazon EMR
- Versão 4.9.2 do Amazon EMR
- Versão 4.9.1 do Amazon EMR
- Versão 4.8.5 do Amazon EMR
- Versão 4.8.4 do Amazon EMR
- Versão 4.8.3 do Amazon EMR
- Versão 4.8.2 do Amazon EMR
- Versão 4.8.1 do Amazon EMR
- Versão 4.8.0 do Amazon EMR
- Versão 4.7.4 do Amazon EMR
- Versão 4.7.3 do Amazon EMR
- Versão 4.7.2 do Amazon EMR
- Versão 4.7.1 do Amazon EMR
- Versão 4.7.0 do Amazon EMR
- Versão 4.6.1 do Amazon EMR
- Versão 4.6.0 do Amazon EMR
- Versão 4.5.0 do Amazon EMR
- Versão 4.4.0 do Amazon EMR
- Versão 4.3.0 do Amazon EMR
- Versão 4.2.0 do Amazon EMR
- Versão 4.1.0 do Amazon EMR
- Versão 4.0.0 do Amazon EMR

Versões de aplicações nas versões 4.x do Amazon EMR

Para obter uma tabela abrangente que liste as versões das aplicações disponíveis em cada versão 4.x do Amazon EMR, abra as versões de aplicações nas versões 4.x do Amazon EMR no seu navegador.

Diferenças nas versões 4.x do Amazon EMR

A documentação para atributos do Amazon EMR no Guia de gerenciamento do Amazon EMR especifica a versão do Amazon EMR na qual um atributo foi disponibilizado, além das diferenças aplicáveis entre atributos do Amazon EMR desde a versão 4.0.0.

Desde a versão 5.0.0 do Amazon EMR, algumas aplicações receberam uma atualização de versão significativa, que alterou detalhes de instalação ou operacionais e outras foram promovidas de aplicações em sandbox a aplicações nativas. Cada tópico desta seção apresenta diferenças significativas e específicas das aplicações quando são usadas as versões 4.x do Amazon EMR.

Tópicos

- Aplicações em sandbox
- Considerações sobre o uso do Hive no Amazon EMR 4.x
- Considerações sobre o uso do Pig no Amazon EMR 4.x

Aplicações em sandbox

Ao serem usadas as versões 4.x do Amazon EMR, algumas aplicações são consideradas aplicações em sandbox. As aplicações em sandbox são versões anteriores da aplicação que disponibilizamos no momento da versão inicial do Amazon EMR devido à demanda. Você pode usar o console ou a API para que o Amazon EMR instale aplicativos sandbox da mesma forma que os aplicativos nativos, mas os aplicativos sandbox têm suporte e documentação limitados. AWS CLI As aplicações em Sandbox se tornaram aplicações nativas, totalmente compatíveis com as versões 5.0.0 e posteriores do Amazon EMR. Veja a seguir as aplicações em sandbox nas versões 4.x do Amazon EMR:

- Oozie
- Presto
- Sqoop
- Zeppelin
- ZooKeeper

Quando você instala aplicativos em sandbox, os nomes dos aplicativos são representados pelo sufixo `-sandbox`. Por exemplo, para instalar a versão sandbox do*`Presto`*, use`Presto-sandbox`. A instalação pode demorar mais do que para um aplicativo totalmente compatível. Os números de versão listados para cada aplicativo nesta seção correspondem à versão da comunidade do aplicativo.

Oozie (versões em sandbox)

O Oozie está disponível como uma aplicação em sandbox desde a versão 4.1.0 do Amazon EMR.

Os exemplos do Oozie não são instalados por padrão usando as versões em sandbox. Para instalar os exemplos, execute o SSH no nó principal do cluster e `install-oozie-examples`.

Informações sobre a versão em sandbox do Oozie

Rótulo de versão do Amazon EMR	Oozie (versão em sandbox)	Componentes instalados com o Oozie-Sandbox
emr-4.9.6	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.9.5	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resour

Rótulo de versão do Amazon EMR	Oozie (versão em sandbox)	Componentes instalados com o Oozie-Sandbox
		cemanager, oozie-client, oozie-server
emr-4.9.4	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.9.3	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Oozie (versão em sandbox)	Componentes instalados com o Oozie-Sandbox
emr-4.9.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.9.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Oozie (versão em sandbox)	Componentes instalados com o Oozie-Sandbox
emr-4.8.5	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.8.4	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Oozie (versão em sandbox)	Componentes instalados com o Oozie-Sandbox
emr-4.8.3	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.8.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Oozie (versão em sandbox)	Componentes instalados com o Oozie-Sandbox
emr-4.8.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.8.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Oozie (versão em sandbox)	Componentes instalados com o Oozie-Sandbox
emr-4.7.4	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.7.3	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Oozie (versão em sandbox)	Componentes instalados com o Oozie-Sandbox
emr-4.7.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.7.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Oozie (versão em sandbox)	Componentes instalados com o Oozie-Sandbox
emr-4.7.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.6.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.6.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Oozie (versão em sandbox)	Componentes instalados com o Oozie-Sandbox
emr-4.5.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.4.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.3.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Oozie (versão em sandbox)	Componentes instalados com o Oozie-Sandbox
emr-4.2.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.1.0	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Presto (versões em sandbox)

O Presto está disponível como aplicação em sandbox desde a versão 4.1.0 do Amazon EMR.

Informações sobre a versão do Presto-Sandbox

Rótulo de versão do Amazon EMR	Presto (versão em sandbox)	Componentes instalados com o Presto-Sandbox
emr-4.9.6	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-

Rótulo de versão do Amazon EMR	Presto (versão em sandbox)	Componentes instalados com o Presto-Sandbox
		kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.9.5	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.9.4	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Presto (versão em sandbox)	Componentes instalados com o Presto-Sandbox
emr-4.9.3	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.9.2	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.9.1	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Presto (versão em sandbox)	Componentes instalados com o Presto-Sandbox
emr-4.8.5	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.8.4	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.8.3	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Presto (versão em sandbox)	Componentes instalados com o Presto-Sandbox
emr-4.8.2	0.152.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.8.1	0.151	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.8.0	0.151	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Presto (versão em sandbox)	Componentes instalados com o Presto-Sandbox
emr-4.7.4	0.148	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.7.3	0.148	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.7.2	0.148	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Presto (versão em sandbox)	Componentes instalados com o Presto-Sandbox
emr-4.7.1	0.147	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.7.0	0.147	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.6.1	0.143	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Presto (versão em sandbox)	Componentes instalados com o Presto-Sandbox
emr-4.6.0	0.143	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker
emr-4.5.0	0.140	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker
emr-4.4.0	0.136	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Presto (versão em sandbox)	Componentes instalados com o Presto-Sandbox
emr-4.3.0	0.130	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker
emr-4.2.0	0.125	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker
emr-4.1.0	0.119	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker

Sqoop (versões em sandbox)

O Sqoop está disponível como uma aplicação em sandbox desde a versão 4.4.0 do Amazon EMR.

Informações sobre a versão do Sqoop-Sandbox

Rótulo de versão do Amazon EMR	Sqoop (versão em sandbox)	Componentes instalados com o Sqoop-Sandbox
emr-4.9.6	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.9.5	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.9.4	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Sqoop (versão em sandbox)	Componentes instalados com o Sqoop-Sandbox
emr-4.9.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.9.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.9.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Sqoop (versão em sandbox)	Componentes instalados com o Sqoop-Sandbox
emr-4.8.5	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.8.4	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.8.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Sqoop (versão em sandbox)	Componentes instalados com o Sqoop-Sandbox
emr-4.8.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.8.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.8.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Sqoop (versão em sandbox)	Componentes instalados com o Sqoop-Sandbox
emr-4.7.4	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.7.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.7.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Sqoop (versão em sandbox)	Componentes instalados com o Sqoop-Sandbox
emr-4.7.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.7.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.6.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Sqoop (versão em sandbox)	Componentes instalados com o Sqoop-Sandbox
emr-4.6.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.5.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, sqoop-client
emr-4.4.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, sqoop-client

Zeppelin (versões em sandbox)

O Zeppelin está disponível como uma aplicação em sandbox desde a versão 4.1.0 do Amazon EMR.

Informações sobre a versão do Zeppelin-Sandbox

Rótulo de versão do Amazon EMR	Zeppelin (versão em sandbox)	Componentes instalados com o Zeppelin-Sandbox
emr-4.9.6	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.9.5	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.9.4	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-

Rótulo de versão do Amazon EMR	Zeppelin (versão em sandbox)	Componentes instalados com o Zeppelin-Sandbox
		on-yarn, spark-yarn-slave, zeppelin-server
emr-4.9.3	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.9.2	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Zeppelin (versão em sandbox)	Componentes instalados com o Zeppelin-Sandbox
emr-4.9.1	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.8.5	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Zeppelin (versão em sandbox)	Componentes instalados com o Zeppelin-Sandbox
emr-4.8.4	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.8.3	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Zeppelin (versão em sandbox)	Componentes instalados com o Zeppelin-Sandbox
emr-4.8.2	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.8.1	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Zeppelin (versão em sandbox)	Componentes instalados com o Zeppelin-Sandbox
emr-4.8.0	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.7.4	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Zeppelin (versão em sandbox)	Componentes instalados com o Zeppelin-Sandbox
emr-4.7.3	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.7.2	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Zeppelin (versão em sandbox)	Componentes instalados com o Zeppelin-Sandbox
emr-4.7.1	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.7.0	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.6.1	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Zeppelin (versão em sandbox)	Componentes instalados com o Zeppelin-Sandbox
emr-4.6.0	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.5.0	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.4.0	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Zeppelin (versão em sandbox)	Componentes instalados com o Zeppelin-Sandbox
emr-4.3.0	0.5.5	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.2.0	0.5.5	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.1.0	0.6.0-SNAPSHOT	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

ZooKeeper (versões sandbox)

O Zookeeper está disponível como uma aplicação em sandbox desde a versão 4.6.0 do Amazon EMR.

ZooKeeper-Informações sobre a versão Sandbox

Rótulo de versão do Amazon EMR	ZooKeeper-Versão Sandbox	Componentes instalados com ZooKeeper -Sandbox
emr-4.9.6	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.9.5	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.9.4	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resour

Rótulo de versão do Amazon EMR	ZooKeeper-Versão Sandbox	Componentes instalados com ZooKeeper -Sandbox
		cemanager, zookeeper-client, zookeeper-server
emr-4.9.3	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.9.2	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.9.1	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper-Versão Sandbox	Componentes instalados com ZooKeeper -Sandbox
emr-4.8.5	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.8.4	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.8.3	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper-Versão Sandbox	Componentes instalados com ZooKeeper -Sandbox
emr-4.8.2	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.8.1	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.8.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper-Versão Sandbox	Componentes instalados com ZooKeeper -Sandbox
emr-4.7.4	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.7.3	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.7.2	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper-Versão Sandbox	Componentes instalados com ZooKeeper -Sandbox
emr-4.7.1	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.7.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.6.1	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper-Versão Sandbox	Componentes instalados com ZooKeeper -Sandbox
emr-4.6.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ht tpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Considerações sobre o uso do Hive no Amazon EMR 4.x

Esta seção discute as diferenças que devem ser consideradas ao ser usada a versão 1.0.0 do Hive nas versões 4.x do Amazon EMR em comparação com o Hive 2.x nas versões 5.x do Amazon EMR.

Transações ACID não são compatíveis

O Hive nas versões 4.x do Amazon EMR não é compatível com transações ACID com dados do Hive armazenados no Amazon S3 ao serem usadas versões 4.x. Se você tentar criar uma tabela transacional no Amazon S3, ocorrerá uma exceção.

Leitura e gravação em tabelas no Amazon S3

O Hive nas versões 4.x do Amazon EMR pode gravar diretamente no Amazon S3 sem o uso de arquivos temporários. Isso melhora a performance, mas uma consequência é que você não pode ler e gravar na mesma tabela no Amazon S3 dentro da mesma instrução do Hive. Uma solução é criar e usar uma tabela temporária no HDFS.

O exemplo a seguir mostra como usar várias instruções do Hive para atualizar uma tabela no Amazon S3. As instruções criam uma tabela temporária no HDFS denominada `tmp` com base em uma tabela no Amazon S3 denominada `my_s3_table`. A tabela no Amazon S3 é atualizada com o conteúdo da tabela temporária.

```
CREATE TEMPORARY TABLE tmp LIKE my_s3_table;
INSERT OVERWRITE TABLE tmp SELECT ....;
INSERT OVERWRITE TABLE my_s3_table SELECT * FROM tmp;
```

Log4j vs. Log4j 2

O Hive nas versões 4.x do Amazon EMR usa o Log4j. O padrão é começar com o Log4j 2 versão 5.0.0. Essas versões podem exigir diferentes configurações de registro. Para obter detalhes, consulte Apache Log4j 2.

MapReduce é o mecanismo de execução padrão

As versões de lançamento 4.x do Hive no Amazon EMR são MapReduce usadas como mecanismo de execução padrão. Desde a versão 5.0.0 do Amazon EMR, o Tez é o padrão, fornecendo melhor performance para a maioria dos fluxos de trabalho.

Autorização do Hive

O Hive nas versões 4.x do Amazon EMR é compatível com a autorização do Hive para o HDFS, mas não para o EMRFS e o Amazon S3. Os clusters do Amazon EMR são executados com a autorização desabilitada por padrão.

Comportamento de mesclagem de arquivos do Hive com o Amazon S3

O Hive nas versões 4.x do Amazon EMR mescla arquivos pequenos no final de um trabalho somente mapa se `hive.merge.mapfiles` é `true`. A mesclagem apenas será acionada se o tamanho médio de saída do trabalho for menor do que a configuração de `hive.merge.smallfiles.avgsize`. O Hive no Amazon EMR tem exatamente o mesmo comportamento se o caminho de saída final está no HDFS. Porém, se o caminho de saída estiver no Amazon S3, o parâmetro `hive.merge.smallfiles.avgsize` será ignorado. Nessa situação, a tarefa de mesclagem sempre será acionada se `hive.merge.mapfiles` estiver definido como `true`.

Considerações sobre o uso do Pig no Amazon EMR 4.x

O Pig versão 0.14.0 é instalado em clusters criados usando versões 4.x do Amazon EMR. O Pig foi atualizado para a versão 0.16.0 no Amazon EMR 5.0.0. As diferenças significativas são abordadas abaixo.

Mecanismo de execução padrão diferente

A versão 0.14.0 do Pig nas versões 4.x do Amazon EMR é usada MapReduce como mecanismo de execução padrão. O Pig 0.16.0 e versões posteriores usam o Apache Tez. Você pode definir

explicitamente a classificação `exectype=mapreduce` de `pig-properties` configuração a ser usada MapReduce.

Funções definidas pelo usuário do Dropped Pig () UDFs

As versões personalizadas UDFs que estavam disponíveis no Pig no Amazon EMR 4.x foram descartadas a partir do Pig 0.16.0. A maioria deles UDFs tem funções equivalentes que você pode usar em vez disso. A tabela a seguir lista funções descartadas UDFs e equivalentes. Para obter mais informações, consulte Funções integradas no site do Apache Pig.

UDF descartada	Função equivalente
FORMAT_DT(dtformat, date)	GetHour(data), GetMinute (data), GetMonth (data), GetSecond (data), GetWeek (data), GetYear (data), GetDay (data)
EXTRACT(string, padrão)	REGEX_EXTRACT_ALL(string, pattern)
REPLACE(string, pattern, replacement)	REPLACE(string, pattern, replacement)
DATE_TIME()	ToDate()
DURATION(dt, dt2)	WeeksBetween(dt, dt2), YearsBetween (dt, dt2), (dt, dt2), SecondsBetween (dt, dt2), MonthsBetween (dt, dt2), MinutesBetween (dt, dt2) HoursBetween
EXTRACT_DT(format, date)	GetHour(data), GetMinute (data), GetMonth (data), GetSecond (data), GetWeek (data), GetYear (data), GetDay (data)
OFFSET_DT(date, duration)	AddDuration(data, duração), SubtractDuration (data, duração)
PERIOD(dt, dt2)	WeeksBetween(dt, dt2), YearsBetween (dt, dt2), (dt, dt2), SecondsBetween (dt, dt2), MonthsBetween (dt, dt2), MinutesBetween (dt, dt2) HoursBetween
CAPITALIZE(string)	UCFIRST(string)

UDF descartada	Função equivalente
CONCAT_WITH()	CONCAT()
INDEX_OF()	INDEXOF()
LAST_INDEX_OF()	LAST_INDEXOF()
SPLIT_ON_REGEX()	STRSPLT()
UNCAPITALIZE()	LCFIRST()

Os itens a seguir UDFs foram eliminados sem equivalente: FORMAT (), LOCAL_DATE (), LOCAL_TIME (), CENTER (), LEFT_PAD (), REPEAT (), REPLACE_ONCE (), RIGHT_PAD (), STRIP (), STRIP_END (), STRIP_START (), SWAP_CASE ().

Comandos Grunt descontinuados

Alguns comandos Grunt foram descontinuados a partir do Pig 0.16.0. A tabela a seguir lista os comandos Grunt no Pig 0.14.0 e os comandos equivalentes na versão atual, quando aplicável.

Comandos Grunt do Pig 0.14.0 e seus equivalentes atuais

Comando Grunt do Pig 0.14.0	Comando Grunt do Pig 0.16.0 e versões posteriores
gato < non-hdfs-path >)	fs -cat < non-hdfs-path >;
cd < non-hdfs-path >;	Não há equivalente
ls < non-hdfs-path >;	fs -ls < non-hdfs-path >;
mover < non-hdfs-path > < non-hdfs-path >;	fs -mv < non-hdfs-path > < >; non-hdfs-path
copiar < non-hdfs-path > < non-hdfs-path >;	fs -cp < non-hdfs-path > < non-hdfs-path >;
copyToLocal < non-hdfs-path ><local-path>;	fs - copyToLocal < non-hdfs-path ><local-path>;
copyFromLocal <local-path>< non-hdfs-path >;	fs - copyFromLocal <local-path>< non-hdfs-path >;

Comando Grunt do Pig 0.14.0	Comando Grunt do Pig 0.16.0 e versões posteriores
mkdir < non-hdfs-path >;	fs -mkdir < >; non-hdfs-path
rm < non-hdfs-path >;	fs -rm -r -SkipTrash < >; non-hdfs-path
rmf < >; non-hdfs-path	fs -rm -r -SkipTrash < >; non-hdfs-path

Capacidade removida para diretórios iniciais não HDFS

O Pig 0.14.0 nas versões 4.x do Amazon EMR tem dois mecanismos para permitir a execução de scripts do Pig por usuários que não sejam o usuário do `hadoop` e não tenham diretórios iniciais. O primeiro mecanismo é um fallback automático que define o diretório de trabalho inicial como o diretório raiz caso o diretório inicial não exista. O segunda é uma propriedade `pig.initial.fs.name` que permite alterar o diretório de trabalho inicial.

Esses mecanismos não estão disponíveis desde a versão 5.0.0 do Amazon EMR e os usuários devem ter um diretório inicial no HDFS. Isso não se aplica ao usuário `hadoop`, pois um diretório inicial é provisionado na execução. Os scripts executados usando etapas JAR do Hadoop assumem como padrão o usuário do Hadoop, a menos que outro usuário seja especificado explicitamente com o uso de `command-runner.jar`.

Versão 4.9.6 do Amazon EMR

Versões 4.9.6 da aplicação

Esta versão inclui os seguintes aplicativos: GangliaHBase, HCatalog, Hadoop, Hive, Hue, Mahout, Oozie-Sandbox, Phoenix, PigPresto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox,, ZooKeeper-Sandboxe.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR

- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-4.9.6	emr-4.9.5	emr-4.9.4	emr-4.9.3
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-

	emr-4.9.6	emr-4.9.5	emr-4.9.4	emr-4.9.3
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

Notas da versão 4.9.6

Esta é uma versão de patch para adicionar a autenticação do AWS Signature Version 4 para solicitações ao Amazon S3. Todas as aplicações e componentes são os mesmos que os da versão anterior do Amazon EMR.

> **Important**
>
> Nesta versão, o Amazon EMR usa o AWS Signature Version 4 exclusivamente para autenticar solicitações para o Amazon S3. Para obter mais informações, consulte Novidades.

Versões de componente 4.9.6

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.3.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.2.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.3.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.

Componente	Versão	Descrição
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.17.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-2	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-2	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-2	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.

Componente	Versão	Descrição
hadoop-httpfs-server	2.7.3-amzn-2	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-2	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-2	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.3-amzn-2	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-2	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.2.2	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.2.2	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.2	HBase cliente de linha de comandos.

Componente	Versão	Descrição
hbase-rest-server	1.2.2	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.2	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	1.0.0-amzn-9	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	1.0.0-amzn-9	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	1.0.0-amzn-9	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	1.0.0-amzn-9	O cliente da linha de comando do Hive.
hive-metastore-server	1.0.0-amzn-9	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	1.0.0-amzn-9	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.7.1-amzn-7	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop

Componente	Versão	Descrição
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.7.0- -amzn-1 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.7.0- -amzn-1 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.157.1	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.157.1	O serviço que executa partes de uma consulta.
pig-client	0.14.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	1.6.3	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	1.6.3	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	1.6.3	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	1.6.3	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.6.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.9	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.9	ZooKeeper cliente de linha de comandos.

Classificações de configuração 4.9.6

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-4.9.6

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.

Classificações	Descrição
hbase-site	Altera os valores no arquivo hbase-site.xml HBase do arquivo.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j	Altere os valores nas HCatalog propriedades HCat log4j da Web.
hcatalog-webhcat-site	Altera os valores no arquivo webhcat-site.xml HCat do HCatalog site da Web.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j	Altere os valores no arquivo hive-exec-log 4j.properties do Hive.
hive-log4j	Alterar os valores no arquivo hive-log4j.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.

Classificações	Descrição
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altera os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.

Classificações	Descrição
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.

Classificações	Descrição
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altera os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.

Classificações	Descrição
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altera os valores no ZooKeeper arquivo log4j.properties.

Versão 4.9.5 do Amazon EMR

Versões 4.9.5 da aplicação

Esta versão inclui os seguintes aplicativos: GangliaHBase, HCatalog, Hadoop, Hive, Hue, Mahout, Oozie-Sandbox, Phoenix, PigPresto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox,, ZooKeeper-Sandboxe.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-4.9.5	emr-4.9.4	emr-4.9.3	emr-4.9.2
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-4.9.5	emr-4.9.4	emr-4.9.3	emr-4.9.2
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

Notas da versão 4.9.5

As notas da versão a seguir incluem informações para a versão 4.9.5 do Amazon EMR. As alterações são referentes à versão 4.9.4.

Data da versão inicial: 29 de agosto de 2018

Alterações, melhorias e problemas resolvidos

- HBase
  - Esta versão aborda uma possível vulnerabilidade de segurança.

Versões de componente 4.9.5

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.3.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.2.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.3.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.

Componente	Versão	Descrição
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.17.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-2	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-2	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-2	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.

Componente	Versão	Descrição
hadoop-httpfs-server	2.7.3-amzn-2	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-2	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-2	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.3-amzn-2	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-2	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.2.2	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.2.2	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.2	HBase cliente de linha de comandos.

Componente	Versão	Descrição
hbase-rest-server	1.2.2	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.2	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	1.0.0-amzn-9	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	1.0.0-amzn-9	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	1.0.0-amzn-9	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	1.0.0-amzn-9	O cliente da linha de comando do Hive.
hive-metastore-server	1.0.0-amzn-9	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	1.0.0-amzn-9	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.7.1-amzn-7	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop

Componente	Versão	Descrição
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.7.0- -amzn-1 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.7.0- -amzn-1 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.157.1	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.157.1	O serviço que executa partes de uma consulta.
pig-client	0.14.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	1.6.3	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	1.6.3	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	1.6.3	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	1.6.3	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.6.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.9	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.9	ZooKeeper cliente de linha de comandos.

Classificações de configuração 4.9.5

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-4.9.5

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.

Classificações	Descrição
hbase-site	Altera os valores no arquivo hbase-site.xml HBase do arquivo.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j	Altere os valores nas HCatalog propriedades HCat log4j da Web.
hcatalog-webhcat-site	Altera os valores no arquivo webhcat-site.xml HCat do HCatalog site da Web.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j	Altere os valores no arquivo hive-exec-log 4j.properties do Hive.
hive-log4j	Alterar os valores no arquivo hive-log4j.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.

Classificações	Descrição
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altera os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.

Classificações	Descrição
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.

Classificações	Descrição
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altera os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.

Classificações	Descrição
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altera os valores no ZooKeeper arquivo log4j.properties.

Versão 4.9.4 do Amazon EMR

Versões 4.9.4 da aplicação

Esta versão inclui os seguintes aplicativos: GangliaHBase, HCatalog, Hadoop, Hive, Hue, Mahout, Oozie-Sandbox, Phoenix, PigPresto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox,, ZooKeeper-Sandboxe.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-4.9.4	emr-4.9.3	emr-4.9.2	emr-4.9.1
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-4.9.4	emr-4.9.3	emr-4.9.2	emr-4.9.1
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

Notas da versão 4.9.4

As notas da versão a seguir incluem informações para a versão 4.9.4 do Amazon EMR. As alterações são referentes à versão 4.9.3.

Data da versão inicial: 29 de março de 2018

Alterações, melhorias e problemas resolvidos

- Atualizado o kernel do Amazon Linux da AMI padrão do Amazon Linux para Amazon EMR para abordar possíveis vulnerabilidades.

Versões de componente 4.9.4

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.3.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.2.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.3.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.

Componente	Versão	Descrição
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.17.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-2	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-2	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-2	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.

Componente	Versão	Descrição
hadoop-httpfs-server	2.7.3-amzn-2	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-2	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-2	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.3-amzn-2	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-2	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.2.2	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.2.2	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.2	HBase cliente de linha de comandos.

Componente	Versão	Descrição
hbase-rest-server	1.2.2	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.2	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	1.0.0-amzn-9	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	1.0.0-amzn-9	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	1.0.0-amzn-9	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	1.0.0-amzn-9	O cliente da linha de comando do Hive.
hive-metastore-server	1.0.0-amzn-9	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	1.0.0-amzn-9	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.7.1-amzn-7	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop

Componente	Versão	Descrição
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.7.0- -amzn-1 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.7.0- -amzn-1 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.157.1	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.157.1	O serviço que executa partes de uma consulta.
pig-client	0.14.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	1.6.3	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	1.6.3	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	1.6.3	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	1.6.3	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.6.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.9	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.9	ZooKeeper cliente de linha de comandos.

Classificações de configuração 4.9.4

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-4.9.4

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.

Classificações	Descrição
hbase-site	Altera os valores no arquivo hbase-site.xml HBase do arquivo.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j	Altere os valores nas HCatalog propriedades HCat log4j da Web.
hcatalog-webhcat-site	Altera os valores no arquivo webhcat-site.xml HCat do HCatalog site da Web.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j	Altere os valores no arquivo hive-exec-log 4j.properties do Hive.
hive-log4j	Alterar os valores no arquivo hive-log4j.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.

Classificações	Descrição
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altera os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.

Classificações	Descrição
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.

Classificações	Descrição
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altera os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.

Classificações	Descrição
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altera os valores no ZooKeeper arquivo log4j.properties.

Versão 4.9.3 do Amazon EMR

Versões 4.9.3 da aplicação

Esta versão inclui os seguintes aplicativos: GangliaHBase, HCatalog, Hadoop, Hive, Hue, Mahout, Oozie-Sandbox, Phoenix, PigPresto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox,, ZooKeeper-Sandboxe.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-4.9.3	emr-4.9.2	emr-4.9.1	emr-4.8.5
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-4.9.3	emr-4.9.2	emr-4.9.1	emr-4.8.5
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

Notas da versão 4.9.3

As notas da versão a seguir incluem informações para a versão 4.9.3 do Amazon EMR. As alterações são referentes à versão 4.9.2 do Amazon EMR.

Data da versão inicial: 22 de janeiro de 2018

Alterações, melhorias e problemas resolvidos

- Atualizado o kernel do Amazon Linux da AMI padrão do Amazon Linux para Amazon EMR para abordar vulnerabilidades associadas à execução especulativa (CVE-2017-5715, CVE-2017-5753 e CVE-2017-5754). Para obter mais informações, consulte https://aws.amazon.com/security/security-bulletins/AWS-2018-013/.

Versões de componente 4.9.3

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.3.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.2.0	Bibliotecas convenientes para o ecossistema do Hadoop.

Componente	Versão	Descrição
emr-kinesis	3.3.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.17.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-2	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-2	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-2	O cliente de linha de comando e biblioteca do HDFS

Componente	Versão	Descrição
hadoop-hdfs-namenode	2.7.3-amzn-2	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-2	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-2	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-2	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.3-amzn-2	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-2	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.2.2	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.

Componente	Versão	Descrição
hbase-region-server	1.2.2	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.2	HBase cliente de linha de comandos.
hbase-rest-server	1.2.2	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.2	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	1.0.0-amzn-9	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	1.0.0-amzn-9	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	1.0.0-amzn-9	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	1.0.0-amzn-9	O cliente da linha de comando do Hive.
hive-metastore-server	1.0.0-amzn-9	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.

Componente	Versão	Descrição
hive-server	1.0.0-amzn-9	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.7.1-amzn-7	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.7.0- -amzn-1 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.7.0- -amzn-1 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.157.1	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.157.1	O serviço que executa partes de uma consulta.

Componente	Versão	Descrição
pig-client	0.14.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	1.6.3	Os clientes da linha de comando do Spark.
spark-history-server	1.6.3	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	1.6.3	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	1.6.3	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.6.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.9	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.

Componente	Versão	Descrição
zookeeper-client	3.4.9	ZooKeeper cliente de linha de comandos.

Classificações de configuração 4.9.3

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-4.9.3

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.

Classificações	Descrição
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.
hbase-site	Altera os valores no arquivo hbase-site.xml HBase do arquivo.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j	Altere os valores nas HCatalog propriedades HCat log4j da Web.
hcatalog-webhcat-site	Altera os valores no arquivo webhcat-site.xml HCat do HCatalog site da Web.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j	Altere os valores no arquivo hive-exec-log4j.properties do Hive.
hive-log4j	Alterar os valores no arquivo hive-log4j.properties do Hive.

Classificações	Descrição
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altera os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.

Classificações	Descrição
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.

Classificações	Descrição
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.

Classificações	Descrição
sqoop-oraoop-site	Altera os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altera os valores no ZooKeeper arquivo log4j.properties.

Versão 4.9.2 do Amazon EMR

Versões 4.9.2 da aplicação

Esta versão inclui os seguintes aplicativos: GangliaHBase, HCatalog, Hadoop, Hive, Hue, Mahout, Oozie-Sandbox, Phoenix, PigPresto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox,, ZooKeeper-Sandboxe.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR

- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-4.9.2	emr-4.9.1	emr-4.8.5	emr-4.8.4
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-4.9.2	emr-4.9.1	emr-4.8.5	emr-4.8.4
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-

	emr-4.9.2	emr-4.9.1	emr-4.8.5	emr-4.8.4
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

Notas da versão 4.9.2

As notas da versão a seguir incluem informações para a versão 4.9.2 do Amazon EMR. As alterações são referentes à versão 4.9.1 do Amazon EMR.

Data do release: 13 de julho de 2017

Pequenas alterações, correções de erros e melhorias foram feitas nesta versão.

Versões de componente 4.9.2

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.3.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.2.0	Bibliotecas convenientes para o ecossistema do Hadoop.

Componente	Versão	Descrição
emr-kinesis	3.3.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.17.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-2	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-2	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-2	O cliente de linha de comando e biblioteca do HDFS

Componente	Versão	Descrição
hadoop-hdfs-namenode	2.7.3-amzn-2	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-2	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-2	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-2	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.3-amzn-2	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-2	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.2.2	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.

Componente	Versão	Descrição
hbase-region-server	1.2.2	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.2	HBase cliente de linha de comandos.
hbase-rest-server	1.2.2	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.2	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	1.0.0-amzn-9	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	1.0.0-amzn-9	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	1.0.0-amzn-9	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	1.0.0-amzn-9	O cliente da linha de comando do Hive.
hive-metastore-server	1.0.0-amzn-9	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.

Componente	Versão	Descrição
hive-server	1.0.0-amzn-9	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.7.1-amzn-7	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.7.0- -amzn-1 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.7.0- -amzn-1 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.157.1	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.157.1	O serviço que executa partes de uma consulta.

Componente	Versão	Descrição
pig-client	0.14.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	1.6.3	Os clientes da linha de comando do Spark.
spark-history-server	1.6.3	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	1.6.3	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	1.6.3	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.6.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.9	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.

Componente	Versão	Descrição
zookeeper-client	3.4.9	ZooKeeper cliente de linha de comandos.

Classificações de configuração 4.9.2

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-4.9.2

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.

Classificações	Descrição
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.
hbase-site	Altera os valores no arquivo hbase-site.xml HBase do arquivo.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j	Altere os valores nas HCatalog propriedades HCat log4j da Web.
hcatalog-webhcat-site	Altera os valores no arquivo webhcat-site.xml HCat do HCatalog site da Web.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j	Altere os valores no arquivo hive-exec-log4j.properties do Hive.
hive-log4j	Alterar os valores no arquivo hive-log4j.properties do Hive.

Classificações	Descrição
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altera os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.

Classificações	Descrição
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.

Classificações	Descrição
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.

Classificações	Descrição
sqoop-oraoop-site	Altera os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altera os valores no ZooKeeper arquivo log4j.properties.

Versão 4.9.1 do Amazon EMR

Versões 4.9.1 da aplicação

Esta versão inclui os seguintes aplicativos: GangliaHBase, HCatalog, Hadoop, Hive, Hue, Mahout, Oozie-Sandbox, Phoenix, PigPresto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox,, ZooKeeper-Sandboxe.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR

- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-4.9.1	emr-4.8.5	emr-4.8.4	emr-4.8.3
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-4.9.1	emr-4.8.5	emr-4.8.4	emr-4.8.3
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-

	emr-4.9.1	emr-4.8.5	emr-4.8.4	emr-4.8.3
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

Notas da versão 4.9.1

As notas da versão a seguir incluem informações para a versão 4.9.1 do Amazon EMR. As alterações são referentes à versão 4.8.4 do Amazon EMR.

Data do release: 10 de abril de 2017

Problemas conhecidos das versões anteriores que foram resolvidos

- Envio para backport do HIVE-9976 e do HIVE-10106
- Corrigido um problema no YARN em que um grande número de nós (mais de 2.000) e de contêineres (mais de 5.000) causavam um out-of-memory erro, por exemplo:`"Exception in thread main java.lang.OutOfMemoryError"`.

Alterações e melhorias

- As versões do Amazon EMR agora são baseadas no Amazon Linux 2017.03. Para obter mais informações, consulte https://aws.amazon.com/amazon-linux-ami/2017.03-release-notes/.
- Removido o Python 2.6 da imagem de base Linux do Amazon EMR. Você pode instalar o Python 2.6 manualmente, se necessário.

Versões de componente 4.9.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`*

começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.2.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.2.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.3.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.15.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas

Componente	Versão	Descrição
		coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-2	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-2	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-2	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-2	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-2	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-2	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.7.3-amzn-2	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-2	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.2.2	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.2.2	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.2	HBase cliente de linha de comandos.
hbase-rest-server	1.2.2	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.2	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	1.0.0-amzn-9	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	1.0.0-amzn-9	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	1.0.0-amzn-9	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	1.0.0-amzn-9	O cliente da linha de comando do Hive.
hive-metastore-server	1.0.0-amzn-9	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	1.0.0-amzn-9	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.7.1-amzn-7	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.7.0- -amzn-1 HBase	As bibliotecas do phoenix para servidor e cliente

Componente	Versão	Descrição
phoenix-query-server	4.7.0- -amzn-1 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.157.1	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.157.1	O serviço que executa partes de uma consulta.
pig-client	0.14.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	1.6.3	Os clientes da linha de comando do Spark.
spark-history-server	1.6.3	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	1.6.3	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	1.6.3	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.

Componente	Versão	Descrição
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.6.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.9	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.9	ZooKeeper cliente de linha de comandos.

Classificações de configuração 4.9.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-4.9.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.

Classificações	Descrição
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altera os valores no arquivo hbase-site.xml HBase do arquivo.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.

Classificações	Descrição
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j	Altere os valores nas HCatalog propriedades HCat log4j da Web.
hcatalog-webhcat-site	Altera os valores no arquivo webhcat-site.xml HCat do HCatalog site da Web.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j	Altere os valores no arquivo hive-exec-log 4j.properties do Hive.
hive-log4j	Alterar os valores no arquivo hive-log4j.propert ies do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.x ml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properti es do Hadoop.

Classificações	Descrição
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altera os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.

Classificações	Descrição
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.

Classificações	Descrição
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altera os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altera os valores no ZooKeeper arquivo log4j.properties.

Versão 4.8.5 do Amazon EMR

Versões 4.8.5 da aplicação

Esta versão inclui os seguintes aplicativos: GangliaHBase, HCatalog, Hadoop, Hive, Hue, Mahout, Oozie-Sandbox, Phoenix, PigPresto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox,, ZooKeeper-Sandboxe.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-4.8.5	emr-4.8.4	emr-4.8.3	emr-4.8.2
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2

	emr-4.8.5	emr-4.8.4	emr-4.8.3	emr-4.8.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.152.3
Spark	1.6.3	1.6.3	1.6.3	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6

	emr-4.8.5	emr-4.8.4	emr-4.8.3	emr-4.8.2
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.8

Notas da versão 4.8.5

Esta é uma versão de patch para adicionar a autenticação do AWS Signature Version 4 para solicitações ao Amazon S3. Todas as aplicações e componentes são os mesmos que os da versão anterior do Amazon EMR.

> **Important**
>
> Nesta versão, o Amazon EMR usa o AWS Signature Version 4 exclusivamente para autenticar solicitações para o Amazon S3. Para obter mais informações, consulte Novidades.

Versões de componente 4.8.5

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *CommunityVersion*-amzn-*EmrVersion*. O *EmrVersion* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado myapp-component com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como 2.2-amzn-2.

Componente	Versão	Descrição
emr-ddb	4.2.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.2.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.2.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.14.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas

Componente	Versão	Descrição
		coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.7.3-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.2.2	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.2.2	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.2	HBase cliente de linha de comandos.
hbase-rest-server	1.2.2	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.2	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	1.0.0-amzn-8	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	1.0.0-amzn-8	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	1.0.0-amzn-8	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	1.0.0-amzn-8	O cliente da linha de comando do Hive.
hive-metastore-server	1.0.0-amzn-8	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	1.0.0-amzn-8	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.7.1-amzn-7	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.7.0- -amzn-1 HBase	As bibliotecas do phoenix para servidor e cliente

Componente	Versão	Descrição
phoenix-query-server	4.7.0- -amzn-1 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.157.1	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.157.1	O serviço que executa partes de uma consulta.
pig-client	0.14.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	1.6.3	Os clientes da linha de comando do Spark.
spark-history-server	1.6.3	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	1.6.3	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	1.6.3	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.

Componente	Versão	Descrição
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.6.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.9	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.9	ZooKeeper cliente de linha de comandos.

Classificações de configuração 4.8.5

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-4.8.5

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.

Classificações	Descrição
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altera os valores no arquivo hbase-site.xml HBase do arquivo.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.

Classificações	Descrição
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j	Altere os valores nas HCatalog propriedades HCat log4j da Web.
hcatalog-webhcat-site	Altera os valores no arquivo webhcat-site.xml HCat do HCatalog site da Web.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j	Altere os valores no arquivo hive-exec-log 4j.properties do Hive.
hive-log4j	Alterar os valores no arquivo hive-log4j.propert ies do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.x ml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properti es do Hadoop.

Classificações	Descrição
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altera os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.

Classificações	Descrição
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.

Classificações	Descrição
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altera os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altera os valores no ZooKeeper arquivo log4j.properties.

Versão 4.8.4 do Amazon EMR

Versões 4.8.4 da aplicação

Esta versão inclui os seguintes aplicativos: GangliaHBase, HCatalog, Hadoop, Hive, Hue, Mahout, Oozie-Sandbox, Phoenix, PigPresto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox,, ZooKeeper-Sandboxe.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-4.8.4	emr-4.8.3	emr-4.8.2	emr-4.8.1
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2

	emr-4.8.4	emr-4.8.3	emr-4.8.2	emr-4.8.1
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.152.3	0.151
Spark	1.6.3	1.6.3	1.6.2	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6

	emr-4.8.4	emr-4.8.3	emr-4.8.2	emr-4.8.1
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.8	3.4.8

Notas da versão 4.8.4

As notas da versão a seguir incluem informações para a versão 4.8.4 do Amazon EMR. As alterações são referentes à versão 4.8.3 do Amazon EMR.

Data da versão: 7 de fevereiro de 2017

Pequenas alterações, correções de erros e melhorias foram feitas nesta versão.

Versões de componente 4.8.4

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado

`myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.2.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.2.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.2.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.14.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.

Componente	Versão	Descrição
hadoop-client	2.7.3-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.

Componente	Versão	Descrição
hadoop-yarn-timeline-server	2.7.3-amzn-1	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.2.2	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.2.2	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.2	HBase cliente de linha de comandos.
hbase-rest-server	1.2.2	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.2	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	1.0.0-amzn-8	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	1.0.0-amzn-8	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	1.0.0-amzn-8	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.

Componente	Versão	Descrição
hive-client	1.0.0-amzn-8	O cliente da linha de comando do Hive.
hive-metastore-server	1.0.0-amzn-8	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	1.0.0-amzn-8	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.7.1-amzn-7	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.54+	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.7.0- -amzn-1 HBase	As bibliotecas do phoenix para servidor e cliente

Componente	Versão	Descrição
phoenix-query-server	4.7.0- -amzn-1 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.157.1	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.157.1	O serviço que executa partes de uma consulta.
pig-client	0.14.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	1.6.3	Os clientes da linha de comando do Spark.
spark-history-server	1.6.3	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	1.6.3	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	1.6.3	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.

Componente	Versão	Descrição
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.25+	O servidor Apache HTTP.
zeppelin-server	0.6.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.9	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.9	ZooKeeper cliente de linha de comandos.

Classificações de configuração 4.8.4

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-4.8.4

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.

Classificações	Descrição
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altera os valores no arquivo hbase-site.xml HBase do arquivo.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.

Classificações	Descrição
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j	Altere os valores nas HCatalog propriedades HCat log4j da Web.
hcatalog-webhcat-site	Altera os valores no arquivo webhcat-site.xml HCat do HCatalog site da Web.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j	Altere os valores no arquivo hive-exec-log 4j.properties do Hive.
hive-log4j	Alterar os valores no arquivo hive-log4j.propert ies do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.x ml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properti es do Hadoop.

Classificações	Descrição
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altera os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.

Classificações	Descrição
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.

Classificações	Descrição
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altera os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altera os valores no ZooKeeper arquivo log4j.properties.

Versão 4.8.3 do Amazon EMR

Versões 4.8.3 da aplicação

Esta versão inclui os seguintes aplicativos: GangliaHBase, HCatalog, Hadoop, Hive, Hue, Mahout, Oozie-Sandbox, Phoenix, PigPresto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox,, ZooKeeper-Sandboxe.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-4.8.3	emr-4.8.2	emr-4.8.1	emr-4.8.0
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2

	emr-4.8.3	emr-4.8.2	emr-4.8.1	emr-4.8.0
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.152.3	0.151	0.151
Spark	1.6.3	1.6.2	1.6.2	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6

	emr-4.8.3	emr-4.8.2	emr-4.8.1	emr-4.8.0
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.8	3.4.8	3.4.8

Notas da versão 4.8.3

As notas da versão a seguir incluem informações para a versão 4.8.3 do Amazon EMR. As alterações são referentes à versão 4.8.2 do Amazon EMR.

Data do release: 29 de dezembro de 2016

Atualizações

- Atualizado para Presto 0.157.1. Para obter mais informações, consulte Notas de release do Presto na documentação do Presto.
- Atualizado para Spark 1.6.3. Para obter mais informações, consulte as Notas de release do Spark na documentação do Apache Spark.
- Atualizado para ZooKeeper 3.9. Para obter mais informações, consulte ZooKeeper Apache ZooKeeper .

Alterações e melhorias

- Adicionado o suporte para o tipo de instância do Amazon EC2 m4.16xlarge no Amazon EMR versão 4.8.3 e posterior, com exceção das versões 5.0.0, 5.0.3 e 5.2.0.

- As versões do Amazon EMR agora são baseadas no Amazon Linux 2016.09. Para obter mais informações, consulte https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/.

Problemas conhecidos das versões anteriores que foram resolvidos

- Corrigido um problema no Hadoop em que o ReplicationMonitor thread podia ficar preso por um longo tempo devido a uma corrida entre a replicação e a exclusão do mesmo arquivo em um grande cluster.
- Corrigido um problema em que ControlledJob #toString falhava com uma exceção de ponteiro nulo (NPE) quando o status do trabalho não era atualizado com êxito.

Versões de componente 4.8.3

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.2.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.2.0	Bibliotecas convenientes para o ecossistema do Hadoop.

Componente	Versão	Descrição
emr-kinesis	3.2.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.13.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.3-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-1	O cliente de linha de comando e biblioteca do HDFS

Componente	Versão	Descrição
hadoop-hdfs-namenode	2.7.3-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.3-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.3-amzn-1	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.2.2	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.

Componente	Versão	Descrição
hbase-region-server	1.2.2	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.2	HBase cliente de linha de comandos.
hbase-rest-server	1.2.2	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.2	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	1.0.0-amzn-8	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	1.0.0-amzn-8	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	1.0.0-amzn-8	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	1.0.0-amzn-8	O cliente da linha de comando do Hive.
hive-metastore-server	1.0.0-amzn-8	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.

Componente	Versão	Descrição
hive-server	1.0.0-amzn-8	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.7.1-amzn-7	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.52	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.7.0- -amzn-1 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.7.0- -amzn-1 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.157.1	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.157.1	O serviço que executa partes de uma consulta.

Componente	Versão	Descrição
pig-client	0.14.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	1.6.3	Os clientes da linha de comando do Spark.
spark-history-server	1.6.3	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	1.6.3	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	1.6.3	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.23	O servidor Apache HTTP.
zeppelin-server	0.6.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.9	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.

Componente	Versão	Descrição
zookeeper-client	3.4.9	ZooKeeper cliente de linha de comandos.

Classificações de configuração 4.8.3

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-4.8.3

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.

Classificações	Descrição
hbase-metrics	Altere os valores no arquivo HBase hadoop-metrics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-policy.xml do.
hbase-site	Altera os valores no arquivo hbase-site.xml HBase do arquivo.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j	Altere os valores nas HCatalog propriedades HCat log4j da Web.
hcatalog-webhcat-site	Altera os valores no arquivo webhcat-site.xml HCat do HCatalog site da Web.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j	Altere os valores no arquivo hive-exec-log4j.properties do Hive.
hive-log4j	Alterar os valores no arquivo hive-log4j.properties do Hive.

Classificações	Descrição
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altera os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.

Classificações	Descrição
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.

Classificações	Descrição
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.

Classificações	Descrição
sqoop-oraoop-site	Altera os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altera os valores no ZooKeeper arquivo log4j.properties.

Versão 4.8.2 do Amazon EMR

Versões 4.8.2 da aplicação

Esta versão inclui os seguintes aplicativos: GangliaHBase, HCatalog, Hadoop, Hive, Hue, Mahout, Oozie-Sandbox, Phoenix, PigPresto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox,, ZooKeeper-Sandboxe.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR

- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-4.8.2	emr-4.8.1	emr-4.8.0	emr-4.7.4
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.1
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-4.8.2	emr-4.8.1	emr-4.8.0	emr-4.7.4
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.152.3	0.151	0.151	0.148
Spark	1.6.2	1.6.2	1.6.2	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.3
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.5.6
ZooKeeper	-	-	-	-

	emr-4.8.2	emr-4.8.1	emr-4.8.0	emr-4.7.4
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

Notas da versão 4.8.2

As notas da versão a seguir incluem informações para a versão 4.8.2 do Amazon EMR. As alterações são referentes à versão 4.8.0 do Amazon EMR.

Data do release: 24 de outubro de 2016

Atualizações

- Atualizado para Hadoop 2.7.3
- Atualizado para Presto 0.152.3, que inclui o suporte para a interface da web do Presto. Você pode acessar a interface da web do Presto no coordenador do Presto usando a porta 8889. Para obter mais informações sobre a interface da web do Presto, consulte Interface da web na documentação do Presto.
- As versões do Amazon EMR agora são baseadas no Amazon Linux 2016.09. Para obter mais informações, consulte https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/.

Versões de componente 4.8.2

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	4.1.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.1.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.2.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.10.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.

Componente	Versão	Descrição
hadoop-client	2.7.3-amzn-0	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.3-amzn-0	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.3-amzn-0	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.3-amzn-0	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.3-amzn-0	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.3-amzn-0	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.7.3-amzn-0	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.3-amzn-0	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-0	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.

Componente	Versão	Descrição
hadoop-yarn-timeline-server	2.7.3-amzn-0	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.2.2	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.2.2	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.2	HBase cliente de linha de comando.
hbase-rest-server	1.2.2	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.2	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	1.0.0-amzn-7	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	1.0.0-amzn-7	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	1.0.0-amzn-7	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.

Componente	Versão	Descrição
hive-client	1.0.0-amzn-7	O cliente da linha de comando do Hive.
hive-metastore-server	1.0.0-amzn-7	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	1.0.0-amzn-7	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.7.1-amzn-7	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.52	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.0- -1,2 HBase	As bibliotecas do phoenix para servidor e cliente

Componente	Versão	Descrição
phoenix-query-server	4.0- -1,2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.152.3	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.152.3	O serviço que executa partes de uma consulta.
pig-client	0.14.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	1.6.2	Os clientes da linha de comando do Spark.
spark-history-server	1.6.2	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	1.6.2	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	1.6.2	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.

Componente	Versão	Descrição
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.23	O servidor Apache HTTP.
zeppelin-server	0.6.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.8	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de linha de comando.

Classificações de configuração 4.8.2

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-4.8.2

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.

Classificações	Descrição
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altera os valores no arquivo hbase-site.xml HBase do arquivo.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.

Classificações	Descrição
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j	Altere os valores nas HCatalog propriedades HCat log4j da Web.
hcatalog-webhcat-site	Altera os valores no arquivo webhcat-site.xml HCat do HCatalog site da Web.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j	Altere os valores no arquivo hive-exec-log 4j.properties do Hive.
hive-log4j	Alterar os valores no arquivo hive-log4j.propert ies do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.x ml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properti es do Hadoop.

Classificações	Descrição
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altera os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.

Classificações	Descrição
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.

Classificações	Descrição
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altera os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altera os valores no ZooKeeper arquivo log4j.properties.

Versão 4.8.1 do Amazon EMR

Versões 4.8.1 da aplicação

Esta versão inclui os seguintes aplicativos: GangliaHBase, HCatalog, Hadoop, Hive, Hue, Mahout, Oozie-Sandbox, Phoenix, PigPresto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox,, ZooKeeper-Sandboxe.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-4.8.1	emr-4.8.0	emr-4.7.4	emr-4.7.3
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.1	1.2.1

	emr-4.8.1	emr-4.8.0	emr-4.7.4	emr-4.7.3
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.151	0.151	0.148	0.148
Spark	1.6.2	1.6.2	1.6.2	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6

	emr-4.8.1	emr-4.8.0	emr-4.7.4	emr-4.7.3
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.3	0.8.3
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.5.6	0.5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

Notas da versão 4.8.1

Versões de componente 4.8.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	3.2.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.1.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.2.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.9.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.

Componente	Versão	Descrição
hadoop-client	2.7.2-amzn-4	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.2-amzn-4	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.2-amzn-4	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.2-amzn-4	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.2-amzn-4	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.2-amzn-4	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.7.2-amzn-4	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.2-amzn-4	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.7.2-amzn-4	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.

Componente	Versão	Descrição
hadoop-yarn-timeline-server	2.7.2-amzn-4	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.2.2	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.2.2	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.2	HBase cliente de linha de comando.
hbase-rest-server	1.2.2	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.2	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	1.0.0-amzn-7	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	1.0.0-amzn-7	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	1.0.0-amzn-7	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.

Componente	Versão	Descrição
hive-client	1.0.0-amzn-7	O cliente da linha de comando do Hive.
hive-metastore-server	1.0.0-amzn-7	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	1.0.0-amzn-7	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.7.1-amzn-7	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.51	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.0- -1,2 HBase	As bibliotecas do phoenix para servidor e cliente

Componente	Versão	Descrição
phoenix-query-server	4.0- -1,2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.151	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.151	O serviço que executa partes de uma consulta.
pig-client	0.14.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	1.6.2	Os clientes da linha de comando do Spark.
spark-history-server	1.6.2	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	1.6.2	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	1.6.2	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.

Componente	Versão	Descrição
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.23	O servidor Apache HTTP.
zeppelin-server	0.6.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.8	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de linha de comando.

Classificações de configuração 4.8.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-4.8.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.

Classificações	Descrição
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altera os valores no arquivo hbase-site.xml HBase do arquivo.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.

Classificações	Descrição
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j	Altere os valores nas HCatalog propriedades HCat log4j da Web.
hcatalog-webhcat-site	Altera os valores no arquivo webhcat-site.xml HCat do HCatalog site da Web.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j	Altere os valores no arquivo hive-exec-log 4j.properties do Hive.
hive-log4j	Alterar os valores no arquivo hive-log4j.propert ies do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.x ml do Hive Server2.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properti es do Hadoop.

Classificações	Descrição
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altera os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.

Classificações	Descrição
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.

Classificações	Descrição
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altera os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altera os valores no ZooKeeper arquivo log4j.properties.

Versão 4.8.0 do Amazon EMR

Versões 4.8.0 da aplicação

Esta versão inclui os seguintes aplicativos: GangliaHBase, HCatalog, Hadoop, Hive, Hue, Mahout, Oozie-Sandbox, Phoenix, PigPresto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox,, ZooKeeper-Sandboxe.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-4.8.0	emr-4.7.4	emr-4.7.3	emr-4.7.2
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.1	1.2.1	1.2.1

	emr-4.8.0	emr-4.7.4	emr-4.7.3	emr-4.7.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.151	0.148	0.148	0.148
Spark	1.6.2	1.6.2	1.6.2	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6

	emr-4.8.0	emr-4.7.4	emr-4.7.3	emr-4.7.2
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.3	0.8.3	0.8.3
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.5.6	0.5.6	0.5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

Notas da versão 4.8.0

As notas da versão a seguir incluem informações para a versão 4.8.0 do Amazon EMR. As alterações são referentes à versão 4.7.2 do Amazon EMR.

Data do release: 7 de setembro de 2016

Atualizações

- Atualizado para 1.2.2 HBase
- Atualizado para Presto-Sandbox 0.151
- Atualizado para Tez 0.8.4
- Atualizado para Zeppelin-Sandbox 0.6.1

Alterações e melhorias

- Corrigido um problema no YARN em que eles ApplicationMaster tentavam limpar contêineres que não existem mais porque suas instâncias foram encerradas.
- Corrigido o URL do hive-server2 para ações do Hive2 em exemplos do Oozie.

- Adicionado o suporte para catálogos Presto adicionais.
- Patches enviados para backport: HIVE-8948, HIVE-12679, HIVE-13405, PHOENIX-3116, HADOOP-12689
- Adicionado o suporte para configurações de segurança, que permitem criar e aplicar opções de criptografia com mais facilidade. Para obter mais informações, consulte Criptografia de dados.

Versões de componente 4.8.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	3.2.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.1.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.2.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.

Componente	Versão	Descrição
emrfs	2.9.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.2-amzn-4	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.2-amzn-4	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.2-amzn-4	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.2-amzn-4	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.2-amzn-4	O endpoint de HTTP para as operações do HDFS.

Componente	Versão	Descrição
hadoop-kms-server	2.7.2-amzn-4	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.2-amzn-4	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.2-amzn-4	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.2-amzn-4	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.2-amzn-4	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.2.2	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.2.2	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.2	HBase cliente de linha de comando.

Componente	Versão	Descrição
hbase-rest-server	1.2.2	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.2	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	1.0.0-amzn-7	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	1.0.0-amzn-7	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	1.0.0-amzn-7	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	1.0.0-amzn-7	O cliente da linha de comando do Hive.
hive-metastore-server	1.0.0-amzn-7	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	1.0.0-amzn-7	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.7.1-amzn-7	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop

Componente	Versão	Descrição
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.51	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.0- -1,2 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.0- -1,2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.151	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.151	O serviço que executa partes de uma consulta.
pig-client	0.14.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	1.6.2	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	1.6.2	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	1.6.2	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	1.6.2	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.4	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.23	O servidor Apache HTTP.
zeppelin-server	0.6.1	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.8	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de linha de comando.

Classificações de configuração 4.8.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-4.8.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.

Classificações	Descrição
hbase-site	Altera os valores no arquivo hbase-site.xml HBase do arquivo.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j	Altere os valores nas HCatalog propriedades HCat log4j da Web.
hcatalog-webhcat-site	Altera os valores no arquivo webhcat-site.xml HCat do HCatalog site da Web.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j	Altere os valores no arquivo hive-exec-log 4j.properties do Hive.
hive-log4j	Alterar os valores no arquivo hive-log4j.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hiveserver2-site	Alterar os valores no arquivo hiveserver2-site.xml do Hive Server2.

Classificações	Descrição
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altera os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.

Classificações	Descrição
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-blackhole	Alterar os valores no arquivo blackhole.properties do Presto.
presto-connector-cassandra	Alterar os valores no arquivo cassandra.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
presto-connector-jmx	Alterar os valores no arquivo jmx.properties do Presto.
presto-connector-kafka	Alterar os valores no arquivo kafka.properties do Presto.
presto-connector-localfile	Alterar os valores no arquivo localfile.properties do Presto.
presto-connector-mongodb	Alterar os valores no arquivo mongodb.properties do Presto.

Classificações	Descrição
presto-connector-mysql	Alterar os valores no arquivo mysql.properties do Presto.
presto-connector-postgresql	Alterar os valores no arquivo postgresql.properties do Presto.
presto-connector-raptor	Alterar os valores no arquivo raptor.properties do Presto.
presto-connector-redis	Alterar os valores no arquivo redis.properties do Presto.
presto-connector-tpch	Alterar os valores no arquivo tpch.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altera os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.

Classificações	Descrição
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altera os valores no ZooKeeper arquivo log4j.properties.

Versão 4.7.4 do Amazon EMR

Versões 4.7.4 da aplicação

Esta versão inclui os seguintes aplicativos: GangliaHBase, HCatalog, Hadoop, Hive, Hue, Mahout, Oozie-Sandbox, Phoenix, PigPresto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox,, ZooKeeper-Sandboxe.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-4.7.4	emr-4.7.3	emr-4.7.2	emr-4.7.1
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.1	1.2.1	1.2.1	1.2.1
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-4.7.4	emr-4.7.3	emr-4.7.2	emr-4.7.1
Mahout	0.12.2	0.12.2	0.12.2	0.12.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.148	0.148	0.148	0.147
Spark	1.6.2	1.6.2	1.6.2	1.6.1
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.3	0.8.3	0.8.3	0.8.3
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

Notas da versão 4.7.4

Esta é uma versão de patch para adicionar a autenticação do AWS Signature Version 4 para solicitações ao Amazon S3. Todas as aplicações e componentes são os mesmos que os da versão anterior do Amazon EMR.

> **Important**
>
> Nesta versão, o Amazon EMR usa o AWS Signature Version 4 exclusivamente para autenticar solicitações para o Amazon S3. Para obter mais informações, consulte Novidades.

Versões de componente 4.7.4

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	3.2.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.1.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.2.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.

Componente	Versão	Descrição
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.8.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.2-amzn-3	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.2-amzn-3	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.2-amzn-3	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.2-amzn-3	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.

Componente	Versão	Descrição
hadoop-httpfs-server	2.7.2-amzn-3	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.2-amzn-3	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.2-amzn-3	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.2-amzn-3	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resourcemanager	2.7.2-amzn-3	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.2-amzn-3	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.2.1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.2.1	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.1	HBase cliente de linha de comando.

Componente	Versão	Descrição
hbase-rest-server	1.2.1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.1	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	1.0.0-amzn-6	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	1.0.0-amzn-6	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	1.0.0-amzn-6	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	1.0.0-amzn-6	O cliente da linha de comando do Hive.
hive-metastore-server	1.0.0-amzn-6	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	1.0.0-amzn-6	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.7.1-amzn-7	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop

Componente	Versão	Descrição
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.46	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.0- -1,2 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.0- -1,2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.148	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.148	O serviço que executa partes de uma consulta.
pig-client	0.14.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	1.6.2	Os clientes da linha de comando do Spark.

Componente	Versão	Descrição
spark-history-server	1.6.2	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	1.6.2	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	1.6.2	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.3	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.23	O servidor Apache HTTP.
zeppelin-server	0.5.6-incubating	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.8	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de linha de comando.

Classificações de configuração 4.7.4

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-4.7.4

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.

Classificações	Descrição
hbase-site	Altera os valores no arquivo hbase-site.xml HBase do arquivo.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j	Altere os valores nas HCatalog propriedades HCat log4j da Web.
hcatalog-webhcat-site	Altera os valores no arquivo webhcat-site.xml HCat do HCatalog site da Web.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j	Altere os valores no arquivo hive-exec-log 4j.properties do Hive.
hive-log4j	Alterar os valores no arquivo hive-log4j.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.

Classificações	Descrição
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altera os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.

Classificações	Descrição
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altera os valores no arquivo oraoop-site.xml OraOop do Sqoop.

Classificações	Descrição
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altera os valores no ZooKeeper arquivo log4j.properties.

Versão 4.7.3 do Amazon EMR

Versões 4.7.3 da aplicação

Esta versão inclui os seguintes aplicativos: GangliaHBase, HCatalog, Hadoop, Hive, Hue, Mahout, Oozie-Sandbox, Phoenix, PigPresto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox,, ZooKeeper-Sandboxe.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-4.7.3	emr-4.7.2	emr-4.7.1	emr-4.7.0
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.1	1.2.1	1.2.1	1.2.1
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-4.7.3	emr-4.7.2	emr-4.7.1	emr-4.7.0
Mahout	0.12.2	0.12.2	0.12.0	0.12.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.148	0.148	0.147	0.147
Spark	1.6.2	1.6.2	1.6.1	1.6.1
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.3	0.8.3	0.8.3	0.8.3
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

Notas da versão 4.7.3

Versões de componente 4.7.3

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	3.2.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.1.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.2.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.8.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do

Componente	Versão	Descrição
		ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.2-amzn-3	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.2-amzn-3	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.2-amzn-3	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.2-amzn-3	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.2-amzn-3	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.2-amzn-3	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er

Componente	Versão	Descrição
hadoop-mapred	2.7.2-amzn-3	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.2-amzn-3	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.7.2-amzn-3	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hadoop-yarn-timeline-server	2.7.2-amzn-3	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.2.1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.2.1	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.1	HBase cliente de linha de comando.
hbase-rest-server	1.2.1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.1	Serviço que fornece um endpoint Thrift para. HBase

Componente	Versão	Descrição
hcatalog-client	1.0.0-amzn-6	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	1.0.0-amzn-6	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	1.0.0-amzn-6	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	1.0.0-amzn-6	O cliente da linha de comando do Hive.
hive-metastore-server	1.0.0-amzn-6	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	1.0.0-amzn-6	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.7.1-amzn-7	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.46	O servidor de banco de dados MySQL.

Componente	Versão	Descrição
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.0- -1,2 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.0- -1,2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.148	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.148	O serviço que executa partes de uma consulta.
pig-client	0.14.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	1.6.2	Os clientes da linha de comando do Spark.
spark-history-server	1.6.2	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.

Componente	Versão	Descrição
spark-on-yarn	1.6.2	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	1.6.2	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.3	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.23	O servidor Apache HTTP.
zeppelin-server	0.5.6-incubating	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.8	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de linha de comando.

Classificações de configuração 4.7.3

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-4.7.3

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altera os valores no arquivo hbase-site.xml HBase do arquivo.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.

Classificações	Descrição
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j	Altere os valores nas HCatalog propriedades HCat log4j da Web.
hcatalog-webhcat-site	Altera os valores no arquivo webhcat-site.xml HCat do HCatalog site da Web.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j	Altere os valores no arquivo hive-exec-log 4j.properties do Hive.
hive-log4j	Alterar os valores no arquivo hive-log4j.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.

Classificações	Descrição
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altera os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.

Classificações	Descrição
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altera os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.

Classificações	Descrição
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altera os valores no ZooKeeper arquivo log4j.properties.

Versão 4.7.2 do Amazon EMR

Versões 4.7.2 da aplicação

Esta versão inclui os seguintes aplicativos: GangliaHBase, HCatalog, Hadoop, Hive, Hue, Mahout, Oozie-Sandbox, Phoenix, PigPresto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox,, ZooKeeper-Sandboxe.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-4.7.2	emr-4.7.1	emr-4.7.0	emr-4.6.1
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.27
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.1	1.2.1	1.2.1	1.2.0
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-4.7.2	emr-4.7.1	emr-4.7.0	emr-4.6.1
Mahout	0.12.2	0.12.0	0.12.0	0.11.1
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.148	0.147	0.147	0.143
Spark	1.6.2	1.6.1	1.6.1	1.6.1
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.3	0.8.3	0.8.3	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

Notas da versão 4.7.2

As notas da versão a seguir incluem informações para a versão 4.7.2 do Amazon EMR.

Data do release: 15 de julho de 2016

Atributos

- Atualizado para Mahout 0.12.2
- Atualizado para Presto 0.148
- Atualizado para Spark 1.6.2
- Agora você pode criar um AWSCredentials provedor para uso com o EMRFS usando um URI como parâmetro. Para obter mais informações, consulte Criar um AWSCredentials provedor para o EMRFS.
- O EMRFS agora permite que os usuários configurem um endpoint personalizado do Dynamo para seus metadados de visualização consistente usando a propriedade `fs.s3.consistent.dynamodb.endpoint` em `emrfs-site.xml`.
- Adicionado um script em `/usr/bin` chamado `spark-example`, que encapsula `/usr/lib/spark/spark/bin/run-example` para que você possa executar exemplos diretamente. Por exemplo, para executar o SparkPi exemplo que vem com a distribuição do Spark, você pode executar a `spark-example SparkPi 100` partir da linha de comando ou usando `command-runner.jar` como uma etapa na API.

Problemas conhecidos das versões anteriores que foram resolvidos

- Corrigido um problema em que o Oozie não tinha o `spark-assembly.jar` no local correto quando o Spark também estava instalado, o que resultava em falha para iniciar aplicativos do Spark com o Oozie.
- Corrigido um problema com o registro baseado em Log4j do Spark em contêineres do YARN.

Versões de componente 4.7.2

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a

versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *CommunityVersion*-amzn-*EmrVersion*. O *EmrVersion* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	3.2.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.1.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.2.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.8.0	O conector do Amazon S3 para aplicações do ecossistema do Hadoop.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.

Componente	Versão	Descrição
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.2-amzn-3	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.2-amzn-3	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.2-amzn-3	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.2-amzn-3	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.2-amzn-3	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.2-amzn-3	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.7.2-amzn-3	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.2-amzn-3	O serviço do YARN para o gerenciamento de contêineres em um nó individual.

Componente	Versão	Descrição
hadoop-yarn-resourcemanager	2.7.2-amzn-3	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuídos.
hadoop-yarn-timeline-server	2.7.2-amzn-3	O serviço para recuperar informações atuais e históricas dos aplicativos do YARN.
hbase-hmaster	1.2.1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.2.1	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.1	HBase cliente de linha de comando.
hbase-rest-server	1.2.1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.1	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	1.0.0-amzn-6	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	1.0.0-amzn-6	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.

Componente	Versão	Descrição
hcatalog-webhcat-server	1.0.0-amzn-6	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	1.0.0-amzn-6	O cliente da linha de comando do Hive.
hive-metastore-server	1.0.0-amzn-6	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	1.0.0-amzn-6	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.7.1-amzn-7	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.2	A biblioteca de Machine Learning.
mysql-server	5.5.46	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.0- -1,2 HBase	As bibliotecas do phoenix para servidor e cliente

Componente	Versão	Descrição
phoenix-query-server	4.0- -1,2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.148	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.148	O serviço que executa partes de uma consulta.
pig-client	0.14.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	1.6.2	Os clientes da linha de comando do Spark.
spark-history-server	1.6.2	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	1.6.2	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	1.6.2	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.

Componente	Versão	Descrição
tez-on-yarn	0.8.3	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.23	O servidor Apache HTTP.
zeppelin-server	0.5.6-incubating	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.8	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de linha de comando.

Classificações de configuração 4.7.2

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-4.7.2

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.

Classificações	Descrição
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hadoop-ssl-server	Alterar a configuração do servidor SSL no Hadoop
hadoop-ssl-client	Alterar a configuração do cliente SSL no Hadoop
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altera os valores no arquivo hbase-site.xml HBase do arquivo.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.

Classificações	Descrição
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j	Altere os valores nas HCatalog propriedades HCat log4j da Web.
hcatalog-webhcat-site	Altera os valores no arquivo webhcat-site.xml HCat do HCatalog site da Web.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j	Altere os valores no arquivo hive-exec-log 4j.properties do Hive.
hive-log4j	Alterar os valores no arquivo hive-log4j.propert ies do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properti es do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.

Classificações	Descrição
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altera os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.

Classificações	Descrição
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altera os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.

Classificações	Descrição
zookeeper-log4j	Altera os valores no ZooKeeper arquivo log4j.properties.

Versão 4.7.1 do Amazon EMR

Versões 4.7.1 da aplicação

Esta versão inclui os seguintes aplicativos: GangliaHBase, HCatalog, Hadoop, Hive, Hue, Mahout, Oozie-Sandbox, Phoenix, PigPresto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox,, ZooKeeper-Sandboxe.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-4.7.1	emr-4.7.0	emr-4.6.1	emr-4.6.0
AWS SDK for Java	1.10.75	1.10.75	1.10.27	1.10.27
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
AmazonCloudWatchAgent	-	-	-	-

	emr-4.7.1	emr-4.7.0	emr-4.6.1	emr-4.6.0
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.1	1.2.1	1.2.0	1.2.0
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.0	0.12.0	0.11.1	0.11.1
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-

	emr-4.7.1	emr-4.7.0	emr-4.6.1	emr-4.6.0
Presto-Sandbox	0.147	0.147	0.143	0.143
Spark	1.6.1	1.6.1	1.6.1	1.6.1
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.3	0.8.3	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

Notas da versão 4.7.1

As notas da versão a seguir incluem informações para a versão 4.7.1 do Amazon EMR.

Data do release: 10 de junho de 2016

Problemas conhecidos das versões anteriores que foram resolvidos

- Corrigido um problema que estendia o tempo de inicialização de clusters iniciados em uma VPC com sub-redes privadas. O bug só afetava os clusters iniciados com a versão 4.7.0. do Amazon EMR.
- Corrigido um problema que tratava incorretamente a listagem de arquivos no Amazon EMR para os clusters iniciados com a versão 4.7.0 do Amazon EMR.

Versões de componente 4.7.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	3.1.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.0.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.2.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.7.1	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop,

Componente	Versão	Descrição
		juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.2-amzn-2	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.2-amzn-2	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.2-amzn-2	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.2-amzn-2	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.2-amzn-2	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.2-amzn-2	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider

Componente	Versão	Descrição
hadoop-mapred	2.7.2-amzn-2	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.2-amzn-2	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.7.2-amzn-2	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hadoop-yarn-timeline-server	2.7.2-amzn-2	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.2.1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.2.1	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.1	HBase cliente de linha de comando.
hbase-rest-server	1.2.1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.1	Serviço que fornece um endpoint Thrift para. HBase

Componente	Versão	Descrição
hcatalog-client	1.0.0-amzn-5	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	1.0.0-amzn-5	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	1.0.0-amzn-5	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	1.0.0-amzn-5	O cliente da linha de comando do Hive.
hive-metastore-server	1.0.0-amzn-5	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	1.0.0-amzn-5	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.7.1-amzn-7	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.0	A biblioteca de Machine Learning.
mysql-server	5.5.46	O servidor de banco de dados MySQL.

Componente	Versão	Descrição
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.0- -1,2 HBase	As bibliotecas do phoenix para servidor e cliente
phoenix-query-server	4.0- -1,2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.147	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.147	O serviço que executa partes de uma consulta.
pig-client	0.14.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	1.6.1	Os clientes da linha de comando do Spark.
spark-history-server	1.6.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.

Componente	Versão	Descrição
spark-on-yarn	1.6.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	1.6.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
tez-on-yarn	0.8.3	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.18	O servidor Apache HTTP.
zeppelin-server	0.5.6-incubating	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.8	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de linha de comando.

Classificações de configuração 4.7.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-4.7.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altera os valores no arquivo hbase-site.xml HBase do arquivo.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.

Classificações	Descrição
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j	Altere os valores nas HCatalog propriedades HCat log4j da Web.
hcatalog-webhcat-site	Altera os valores no arquivo webhcat-site.xml HCat do HCatalog site da Web.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j	Altere os valores no arquivo hive-exec-log 4j.properties do Hive.
hive-log4j	Alterar os valores no arquivo hive-log4j.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.

Classificações	Descrição
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altera os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.

Classificações	Descrição
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altera os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.

Classificações	Descrição
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altera os valores no ZooKeeper arquivo log4j.properties.

Versão 4.7.0 do Amazon EMR

Versões 4.7.0 da aplicação

Esta versão inclui os seguintes aplicativos: GangliaHBase, HCatalog, Hadoop, Hive, Hue, Mahout, Oozie-Sandbox, Phoenix, PigPresto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox,, ZooKeeper-Sandboxe.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-4.7.0	emr-4.6.1	emr-4.6.0	emr-4.5.0
AWS SDK for Java	1.10.75	1.10.27	1.10.27	1.10.27
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas

	emr-4.7.0	emr-4.6.1	emr-4.6.0	emr-4.5.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.1	1.2.0	1.2.0	-
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.7.7.7.7.7.	2.7.7.7.7.7.7.	2.7.7.7.7.7.7.	2.7.7.7.7.7.7.
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.0	0.11.1	0.11.1	0.11.1
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	-	-	-

	emr-4.7.0	emr-4.6.1	emr-4.6.0	emr-4.5.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.147	0.143	0.143	0.140
Spark	1.6.1	1.6.1	1.6.1	1.6.1
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.3	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	-

Notas da versão 4.7.0

> **Important**
>
> A versão 4.7.0 do Amazon EMR foi descontinuada. Use a versão 4.7.1 ou posterior do Amazon EMR.

Data do release: 2 de junho de 2016

Atributos

- Adicionado o Apache Phoenix 4.7.0
- Adicionado o Apache Tez 0.8.3
- Atualizado para 1. HBase
- Atualizado para Mahout 0.12.0
- Atualizado para Presto 0.147
- Atualizado AWS SDK para Java para 1.10.75
- O sinalizador final foi removido da propriedade `mapreduce.cluster.local.dir` em `mapred-site.xml` para permitir que os usuários executem o Pig no modo local.
- Drivers JDBC do Amazon Redshift disponíveis no cluster

  Os drivers JDBC do Amazon Redshift agora estão incluídos em `/usr/share/aws/redshift/jdbc`. `/usr/share/aws/redshift/jdbc/RedshiftJDBC41.jar` é o driver do Amazon Redshift compatível com JDBC 4.1 e `/usr/share/aws/redshift/jdbc/RedshiftJDBC4.jar` é o driver do Amazon Redshift compatível com JDBC 4.0. Para obter mais informações, consulte Configurar uma conexão JDBC no Guia de gerenciamento do Amazon Redshift.
- Java 8

  Com exceção do Presto, o OpenJDK 1.7 é o JDK padrão usado para todos os aplicativos. No entanto, o OpenJDK 1.7 e o 1.8 estão instalados. Para obter mais informações sobre como configurar `JAVA_HOME` para aplicativos, consulte Configuração de aplicativos para usar Java 8.

Problemas conhecidos das versões anteriores que foram resolvidos

- Corrigido um problema do kernel que afetava significativamente a performance dos volumes do EBS no disco rígido com throughput otimizado (st1) para o Amazon EMR no emr-4.6.0.
- Corrigido um problema em que havia uma falha do cluster se qualquer zona de criptografia do HDFS fosse especificada sem escolher o Hadoop como um aplicativo.
- Alterada a política de gravação padrão do HDFS de `RoundRobin` para `AvailableSpaceVolumeChoosingPolicy`. Alguns volumes não foram utilizados adequadamente com a RoundRobin configuração, o que resultou em falhas nos nós principais e em um HDFS não confiável.

- Corrigido um problema com a CLI do EMRFS, que causava uma exceção ao criar a tabela de metadados do DynamoDB padrão para visualizações consistentes.
- Corrigido um problema de deadlock no EMRFS que potencialmente ocorria durante a renomeação do multipart e operações de cópia.
- Corrigido um problema com o EMRFS que fazia com que o CopyPart tamanho padrão fosse 5 MB. O padrão agora está definido corretamente como 128 MB.
- Corrigido um problema com a configuração de inicialização do Zeppelin que potencialmente impedia a interrupção do serviço.
- Corrigido um problema com o Spark e o Zeppelin que impedia o uso do esquema de URI `s3a://` porque `/usr/lib/hadoop/hadoop-aws.jar` não era carregado adequadamente em seus respectivos classpath.
- HUE-2484 enviado para backport.
- Reportou um commit do Hue 3.9.0 (não existe JIRA) para corrigir um problema com a amostra do navegador. HBase
- HIVE-9073 enviado para backport.

Versões de componente 4.7.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	3.1.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.0.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.2.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.4.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.7.1	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.

Componente	Versão	Descrição
hadoop-client	2.7.2-amzn-2	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.2-amzn-2	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.2-amzn-2	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.2-amzn-2	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.2-amzn-2	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.2-amzn-2	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.7.2-amzn-2	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.2-amzn-2	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.7.2-amzn-2	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.

Componente	Versão	Descrição
hadoop-yarn-timeline-server	2.7.2-amzn-2	O serviço para recuperar informações atuais e histórica s dos aplicativos do YARN.
hbase-hmaster	1.2.1	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.2.1	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.1	HBase cliente de linha de comando.
hbase-rest-server	1.2.1	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.1	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	1.0.0-amzn-5	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	1.0.0-amzn-5	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativ os distribuídos.
hcatalog-webhcat-server	1.0.0-amzn-5	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.

Componente	Versão	Descrição
hive-client	1.0.0-amzn-5	O cliente da linha de comando do Hive.
hive-metastore-server	1.0.0-amzn-5	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	1.0.0-amzn-5	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.7.1-amzn-7	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.12.0	A biblioteca de Machine Learning.
mysql-server	5.5.46	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
phoenix-library	4.0- -1,2 HBase	As bibliotecas do phoenix para servidor e cliente

Componente	Versão	Descrição
phoenix-query-server	4.0- -1,2 HBase	Um servidor leve que fornece acesso à JDBC, bem como buffers de protocolo e acesso no formato JSON à API do Avatica
presto-coordinator	0.147	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.147	O serviço que executa partes de uma consulta.
pig-client	0.14.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	1.6.1	Os clientes da linha de comando do Spark.
spark-history-server	1.6.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	1.6.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	1.6.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.

Componente	Versão	Descrição
tez-on-yarn	0.8.3	O aplicativo e as bibliotecas do YARN para Tez.
webserver	2.4.18	O servidor Apache HTTP.
zeppelin-server	0.5.6-incubating	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.8	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de linha de comando.

Classificações de configuração 4.7.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-4.7.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.

Classificações	Descrição
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altera os valores no arquivo hbase-site.xml HBase do arquivo.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j	Altere os valores nas HCatalog propriedades HCat log4j da Web.

Classificações	Descrição
hcatalog-webhcat-site	Altera os valores no arquivo webhcat-site.xml HCat do HCatalog site da Web.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j	Altere os valores no arquivo hive-exec-log 4j.properties do Hive.
hive-log4j	Alterar os valores no arquivo hive-log4j.propert ies do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properti es do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altera os valores no arquivo mapred-site.xml do MapReduce aplicativo.

Classificações	Descrição
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
phoenix-hbase-metrics	Alterar os valores no arquivo hadoop-metrics2-hbase.properties do Phoenix.
phoenix-hbase-site	Alterar os valores no arquivo hbase-site.xml do Phoenix.
phoenix-log4j	Alterar os valores no arquivo log4j.properties do Phoenix.
phoenix-metrics	Alterar os valores no arquivo hadoop-metrics2-phoenix.properties do Phoenix.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.

Classificações	Descrição
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altera os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
tez-site	Alterar os valores no arquivo tez-site.xml do Tez.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altera os valores no ZooKeeper arquivo log4j.properties.

Versão 4.6.1 do Amazon EMR

Versões 4.6.1 da aplicação

Essa versão inclui os seguintes aplicativos: GangliaHBase, HCatalog, Hadoop, Hive, Hue, Mahout, Oozie-Sandbox, PigPresto-Sandbox, Spark, Sqoop-Sandbox, Zeppelin-Sandbox,, ZooKeeper-Sandboxe.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-4.6.1	emr-4.6.0	emr-4.5.0	emr-4.4.0
AWS SDK for Java	1.10.27	1.10.27	1.10.27	1.10.27
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.0	1.2.0	-	-

	emr-4.6.1	emr-4.6.0	emr-4.5.0	emr-4.4.0
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.7.7.7.7.7.	2.7.7.7.7.7.7.	2.7.7.7.7.7.7.	2.7.1
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.1	0.11.1	0.11.1	0.11.1
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.143	0.143	0.140	0.136
Spark	1.6.1	1.6.1	1.6.1	1.6.0
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6

	emr-4.6.1	emr-4.6.0	emr-4.5.0	emr-4.4.0
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	-	-

Notas da versão 4.6.1

Versões de componente 4.6.1

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	3.0.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.0.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.1.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.3.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.6.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.

Componente	Versão	Descrição
hadoop-client	2.7.2-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.2-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.2-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.2-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.2-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.2-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.7.2-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.2-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.7.2-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.

Componente	Versão	Descrição
hbase-hmaster	1.2.0	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.2.0	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.0	HBase cliente de linha de comando.
hbase-rest-server	1.2.0	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.0	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	1.0.0-amzn-4	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	1.0.0-amzn-4	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	1.0.0-amzn-4	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	1.0.0-amzn-4	O cliente da linha de comando do Hive.

Componente	Versão	Descrição
hive-metastore-server	1.0.0-amzn-4	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	1.0.0-amzn-4	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.7.1-amzn-6	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.11.1	A biblioteca de Machine Learning.
mysql-server	5.5	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
presto-coordinator	0.143	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.143	O serviço que executa partes de uma consulta.
pig-client	0.14.0-amzn-0	O cliente da linha de comando do Pig.

Componente	Versão	Descrição
spark-client	1.6.1	Os clientes da linha de comando do Spark.
spark-history-server	1.6.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	1.6.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	1.6.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
webserver	2.4	O servidor Apache HTTP.
zeppelin-server	0.5.6-incubating	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.8	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de linha de comando.

Classificações de configuração 4.6.1

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-4.6.1

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altera os valores no arquivo hbase-site.xml HBase do arquivo.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.

Classificações	Descrição
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j	Altere os valores nas HCatalog propriedades HCat log4j da Web.
hcatalog-webhcat-site	Altera os valores no arquivo webhcat-site.xml HCat do HCatalog site da Web.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j	Altere os valores no arquivo hive-exec-log 4j.properties do Hive.
hive-log4j	Alterar os valores no arquivo hive-log4j.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.

Classificações	Descrição
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altera os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.

Classificações	Descrição
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altera os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altera os valores no ZooKeeper arquivo log4j.properties.

Versão 4.6.0 do Amazon EMR

Versões 4.6.0 da aplicação

Essa versão inclui os seguintes aplicativos: GangliaHBase, HCatalog, Hadoop, Hive, Hue, Mahout, Oozie-Sandbox, PigPresto-Sandbox, Spark, Sqoop-Sandbox, Zeppelin-Sandbox,, ZooKeeper-Sandboxe.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-4.6.0	emr-4.5.0	emr-4.4.0	emr-4.3.0
AWS SDK for Java	1.10.27	1.10.27	1.10.27	1.10.27
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.0	-	-	-

	emr-4.6.0	emr-4.5.0	emr-4.4.0	emr-4.3.0
HCatalog	1.0.0	1.0.0	1.0.0	-
Hadoop	2.7.7.7.7.7.7.	2.7.7.7.7.7.7.	2.7.1	2.7.1
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.1	0.11.1	0.11.1	0.11.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.143	0.140	0.136	0.130
Spark	1.6.1	1.6.1	1.6.0	1.6.0
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	-

	emr-4.6.0	emr-4.5.0	emr-4.4.0	emr-4.3.0
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.5
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	-	-	-

Notas da versão 4.6.0

As notas da versão a seguir incluem informações para a versão 4.6.0 do Amazon EMR.

- adicionada à HBase versão 1.0
- Adicionado o ZooKeeper-Sandbox 3.4.8
- Atualizado para Presto-Sandbox 0.143
- As versões do Amazon EMR agora são baseadas no Amazon Linux 2016.03.0. Para obter mais informações, consulte https://aws.amazon.com/amazon-linux-ami/2016.03-release-notes/.
- Problema afetando os tipos de volumes do EBS no disco rígido com throughput otimizado (st1)

  Um problema no kernel do Linux versões 4.2 e superior afeta significativamente o desempenho nos volumes do EBS no disco rígido com throughput otimizado (st1) para o EMR. Esta versão (emr-4.6.0) usa uma versão do kernel 4.4.5 e, portanto, é afetada. Por isso, recomendamos não usar o emr-4.6.0 se você deseja usar volumes do EBS no st1. Você pode usar emr-4.5.0 ou versões anteriores do Amazon EMR com o st1 sem impacto algum. Além disso, fornecemos a correção com futuras versões.
- Padrões Python

O Python 3.4 agora está instalado por padrão, mas o Python 2.7 permanece como o sistema padrão. Você pode configurar o Python 3.4 como o padrão do sistema usando uma ação de bootstrap; você pode usar a API de configuração para definir a exportação de PYSPARK_PYTHON na classificação `spark-env` para afetar a versão `/usr/bin/python3.4` do Python usada por. PySpark

- Java 8

  Com exceção do Presto, o OpenJDK 1.7 é o JDK padrão usado para todos os aplicativos. No entanto, o OpenJDK 1.7 e o 1.8 estão instalados. Para obter mais informações sobre como configurar `JAVA_HOME` para aplicativos, consulte Configuração de aplicativos para usar Java 8.

Problemas conhecidos das versões anteriores que foram resolvidos

- Corrigido um problema em que, às vezes, ocorria uma falha aleatória no provisionamento de aplicativos devido a uma senha gerada.
- Anteriormente, `mysqld` estava instalado em todos os nós. Agora, ele só está instalado na instância principal e somente se o aplicativo escolhido incluir `mysql-server` como componente. Atualmente, os seguintes aplicativos incluem o `mysql-server` componente: Hive HCatalog, Hue, Presto-Sandbox e Sqoop-Sandbox.
- Alterado o `yarn.scheduler.maximum-allocation-vcores` para 80 do padrão de 32, que corrige um problema apresentado na emr-4.4.0 que ocorre, principalmente, com o Spark enquanto usa a opção `maximizeResourceAllocation` em um cluster cujo tipo de instância core é um dos grandes tipos de instância que têm os vcores do YARN definidos como maior do que 32, ou seja, c4.8xlarge, cc2.8xlarge, hs1.8xlarge, i2.8xlarge, m2.4xlarge, r3.8xlarge, d2.8xlarge ou m4.10xlarge foram afetados por esse problema.
- s3-dist-cp agora usa o EMRFS para todas as indicações do Amazon S3 e não mais estágios para um diretório HDFS temporário.
- Corrigido um problema com o tratamento de exceções para os multipart uploads de criptografia no lado do cliente.
- Adicionada uma opção para permitir que os usuários alterem a classe de armazenamento do Amazon S3. Por padrão, essa configuração é `STANDARD`. A configuração da classificação de configuração `emrfs-site` é `fs.s3.storageClass` e os valores possíveis são `STANDARD`, `STANDARD_IA`e `REDUCED_REDUNDANCY`. Para obter mais informações sobre classes de armazenamento, consulte Classes de armazenamento no Guia do usuário do Amazon Simple Storage Service.

Versões de componente 4.6.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	3.0.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.0.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.1.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.3.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.6.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop,

Componente	Versão	Descrição
		juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.7.2-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.2-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.2-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.2-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.2-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.2-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider

Componente	Versão	Descrição
hadoop-mapred	2.7.2-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.2-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.7.2-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hbase-hmaster	1.2.0	Serviço para um HBase cluster responsável pela coordenação das regiões e execução de comandos administrativos.
hbase-region-server	1.2.0	Serviço para atender uma ou mais HBase regiões.
hbase-client	1.2.0	HBase cliente de linha de comando.
hbase-rest-server	1.2.0	Serviço que fornece um endpoint RESTful HTTP para HBase.
hbase-thrift-server	1.2.0	Serviço que fornece um endpoint Thrift para. HBase
hcatalog-client	1.0.0-amzn-4	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.

Componente	Versão	Descrição
hcatalog-server	1.0.0-amzn-4	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	1.0.0-amzn-4	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	1.0.0-amzn-4	O cliente da linha de comando do Hive.
hive-metastore-server	1.0.0-amzn-4	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	1.0.0-amzn-4	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.7.1-amzn-6	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.11.1	A biblioteca de Machine Learning.
mysql-server	5.5	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.

Componente	Versão	Descrição
oozie-server	4.2.0	O serviço que aceita solicitações de fluxo de trabalho do Oozie.
presto-coordinator	0.143	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.143	O serviço que executa partes de uma consulta.
pig-client	0.14.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	1.6.1	Os clientes da linha de comando do Spark.
spark-history-server	1.6.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	1.6.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	1.6.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.
webserver	2.4	O servidor Apache HTTP.

Componente	Versão	Descrição
zeppelin-server	0.5.6-incubating	O notebook baseado na web que permite um data analytics interativo.
zookeeper-server	3.4.8	O serviço centralizado de manutenção de informações de configuração, nomenclatura, fornecimento de sincronização distribuída, e fornecimento de serviços de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de linha de comando.

Classificações de configuração 4.6.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-4.6.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.

Classificações	Descrição
hbase-env	Mude os valores no HBase ambiente.
hbase-log4j	Altere os valores no arquivo HBase hbase-log 4j.properties.
hbase-metrics	Altere os valores no arquivo HBase hadoop-me trics2-hbaase.properties do.
hbase-policy	Altere os valores no HBase arquivo hbase-pol icy.xml do.
hbase-site	Altera os valores no arquivo hbase-site.xml HBase do arquivo.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j	Altere os valores nas HCatalog propriedades HCat log4j da Web.
hcatalog-webhcat-site	Altera os valores no arquivo webhcat-site.xml HCat do HCatalog site da Web.
hive-env	Alterar os valores no ambiente do Hive.

Classificações	Descrição
hive-exec-log4j	Altere os valores no arquivo hive-exec-log 4j.properties do Hive.
hive-log4j	Alterar os valores no arquivo hive-log4j.propert ies do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properti es do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altera os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.proper ties do Oozie.

Classificações	Descrição
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altera os valores no arquivo oraoop-site.xml OraOop do Sqoop.

Classificações	Descrição
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.
zookeeper-config	Altere os valores no ZooKeeper arquivo zoo.cfg do.
zookeeper-log4j	Altera os valores no ZooKeeper arquivo log4j.properties.

Versão 4.5.0 do Amazon EMR

Versões 4.5.0 da aplicação

Essa versão inclui os seguintes aplicativos: GangliaHCatalog, Hadoop, Hive, Hue, Mahout, Oozie-SandboxPig, Presto-Sandbox, Spark, Sqoop-Sandbox,, Zeppelin-Sandboxe.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-4.5.0	emr-4.4.0	emr-4.3.0	emr-4.2.0
AWS SDK for Java	1.10.27	1.10.27	1.10.27	1.10.27
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.6.0
HBase	-	-	-	-
HCatalog	1.0.0	1.0.0	-	-
Hadoop	2.7.7.7.7.7.7.	2.7.1	2.7.1	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-4.5.0	emr-4.4.0	emr-4.3.0	emr-4.2.0
Mahout	0.11.1	0.11.1	0.11.0	0.11.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.140	0.136	0.130	0.125
Spark	1.6.1	1.6.0	1.6.0	1.5.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.5	0.5.5
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

Notas da versão 4.5.0

As notas da versão a seguir incluem informações para a versão 4.5.0 do Amazon EMR.

Data do release: 4 de abril de 2016

Atributos

- Atualizado para Spark 1.6.1
- Atualizado para Hadoop 2.7.2
- Atualizado para Presto 0.140
- Adicionado AWS KMS suporte do servidor do servidor do servidor do servidor do servidor do servidor do servidor do servidor do servidor do servidor do servidor do servidor do servidor do servidor do servidor do servidor

Problemas conhecidos das versões anteriores que foram resolvidos

- Corrigido um problema em que os servidores MySQL e Apache não iniciavam depois que um nó fosse reinicializado.
- Corrigido um problema em que IMPORT não funcionava corretamente com tabelas não particionadas armazenadas no Amazon S3
- Corrigido um problema em que o Presto exigia que o diretório de preparo fosse `/mnt/tmp` em vez de `/tmp` ao gravar em tabelas do Hive.

Versões de componente 4.5.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado

`myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	3.0.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.0.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.1.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.2.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.5.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.

Componente	Versão	Descrição
hadoop-client	2.7.2-amzn-0	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.2-amzn-0	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.2-amzn-0	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.2-amzn-0	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.2-amzn-0	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.2-amzn-0	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.7.2-amzn-0	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.2-amzn-0	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.7.2-amzn-0	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.

Componente	Versão	Descrição
hcatalog-client	1.0.0-amzn-4	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	1.0.0-amzn-4	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	1.0.0-amzn-4	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	1.0.0-amzn-4	O cliente da linha de comando do Hive.
hive-metastore-server	1.0.0-amzn-4	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	1.0.0-amzn-4	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.7.1-amzn-5	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.11.1	A biblioteca de Machine Learning.
mysql-server	5.5	O servidor de banco de dados MySQL.

Componente	Versão	Descrição
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitações de fluxo de trabalho do Oozie.
presto-coordinator	0.140	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.140	O serviço que executa partes de uma consulta.
pig-client	0.14.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	1.6.1	Os clientes da linha de comando do Spark.
spark-history-server	1.6.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	1.6.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	1.6.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.

Componente	Versão	Descrição
webserver	2.4	O servidor Apache HTTP.
zeppelin-server	0.5.6-incubating	O notebook baseado na web que permite um data analytics interativo.

Classificações de configuração 4.5.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-4.5.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.

Classificações	Descrição
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j	Altere os valores nas HCatalog propriedades HCat log4j da Web.
hcatalog-webhcat-site	Altera os valores no arquivo webhcat-site.xml HCat do HCatalog site da Web.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j	Altere os valores no arquivo hive-exec-log 4j.properties do Hive.
hive-log4j	Alterar os valores no arquivo hive-log4j.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.

Classificações	Descrição
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altera os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.

Classificações	Descrição
spark-env	Alterar os valores no ambiente do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altera os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.

Versão 4.4.0 do Amazon EMR

Versões 4.4.0 da aplicação

Essa versão inclui os seguintes aplicativos: GangliaHCatalog, Hadoop, Hive, Hue, Mahout, Oozie-SandboxPig, Presto-Sandbox, Spark, Sqoop-Sandbox,, Zeppelin-Sandboxe.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR

- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-4.4.0	emr-4.3.0	emr-4.2.0	emr-4.1.0
AWS SDK for Java	1.10.27	1.10.27	1.10.27	Não rastreadas
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.6.0	-
HBase	-	-	-	-
HCatalog	1.0.0	-	-	-
Hadoop	2.7.1	2.7.1	2.6.0	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-

	emr-4.4.0	emr-4.3.0	emr-4.2.0	emr-4.1.0
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.1	0.11.0	0.11.0	0.11.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.0.1
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.136	0.130	0.125	0.119
Spark	1.6.0	1.6.0	1.5.2	1.5.0
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	-	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.5	0.5.5	0.6.0-SNAPSHOT
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

Notas da versão 4.4.0

As notas da versão a seguir incluem informações para a versão 4.4.0 do Amazon EMR.

Data do release: 14 de março de 2016

Atributos

- adicionada à HCatalog versão 1.0
- Adicionado o Sqoop-Sandbox 1.4.6
- Atualizado para Presto 0.136
- Atualizado para Zeppelin 0.5.6
- Atualizado para Mahout 0.11.1
- Habilitada a `dynamicResourceAllocation` por padrão.
- Adicionada uma tabela de todas as classificações de configuração para a versão. Para obter mais informações, consulte a tabela de classificações de configuração em Configurar aplicativos.

Problemas conhecidos das versões anteriores que foram resolvidos

- Corrigido um problema em que a `maximizeResourceAllocation` configuração não reservava memória suficiente para os ApplicationMaster daemons do YARN.
- Corrigido um problema encontrado com um DNS personalizado. Se alguma entrada em `resolve.conf` preceder as entradas personalizadas fornecidas, as entradas personalizadas não serão resolvíveis. Esse comportamento foi afetado por clusters em uma VPC em que o servidor de nomes padrão da VPC está inserido como a entrada superior em `resolve.conf`.
- Corrigido um problema em que o Python padrão mudou para a versão 2.7 e boto não estava instalado para essa versão.
- Corrigido um problema em que contêineres do YARN e aplicativos do Spark geravam um arquivo (rrd) exclusivo do banco de dados Ganglia round robin, que resultava no enchimento do primeiro disco anexado à instância. Devido a essa correção, as métricas no nível do contêiner do YARN e as métricas no nível do aplicativo do Spark foram desativadas.
- Corrigido um problema no log pusher em que ele excluía todas as pastas de log vazias. O efeito era que a CLI do Hive não podia criar logs porque o log pusher estava removendo a pasta `user` vazia em `/var/log/hive`.
- Corrigido um problema com as importações do Hive, que afetava o particionamento e resultava em um erro durante a importação.

- Corrigido um problema em que EMRFS e s3-dist-cp não lidavam adequadamente com nomes de bucket que continham pontos.
- Alterado um comportamento no EMRFS, de modo que nos buckets habilitados para versionamento, o arquivo marcador `_$folder$` não é criado continuamente, o que pode contribuir para um melhor desempenho dos buckets habilitados para versionamento.
- Alterado o comportamento no EMRFS de forma que ele não use os arquivos de instrução, exceto para casos em que a criptografia no lado do cliente esteja habilitada. Se desejar excluir os arquivos de instrução enquanto usa a criptografia no lado do cliente, você poderá definir a propriedade emrfs-site.xml `fs.s3.cse.cryptoStorageMode.deleteInstructionFiles.enabled`, como verdadeiro.
- Alterada a agregação de logs do YARN para reter os logs no destino de agregação por dois dias. O destino padrão é o armazenamento HDFS do cluster. Se desejar mudar essa duração, altere o valor de `yarn.log-aggregation.retain-seconds` usando a classificação de configuração `yarn-site` quando criar seu cluster. Como sempre, você pode salvar os logs de aplicações no Amazon S3 usando o parâmetro `log-uri` quando criar o cluster.

Patches aplicados

- HIVE-9655
- HIVE-9183
- HADOOP-12810

Versões de componente 4.4.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	3.0.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.0.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.1.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.2.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.4.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.

Componente	Versão	Descrição
hadoop-client	2.7.1-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.1-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.1-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.1-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.1-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.1-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.7.1-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.1-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.7.1-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.

Componente	Versão	Descrição
hcatalog-client	1.0.0-amzn-3	O cliente da linha de comando 'hcat' para manipular o hcatalog-server.
hcatalog-server	1.0.0-amzn-3	Fornecimento de serviços HCatalog, uma camada de gerenciamento de tabelas e armazenamento para aplicativos distribuídos.
hcatalog-webhcat-server	1.0.0-amzn-3	Ponto de extremidade HTTP fornecendo uma interface REST para HCatalog.
hive-client	1.0.0-amzn-3	O cliente da linha de comando do Hive.
hive-metastore-server	1.0.0-amzn-3	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	1.0.0-amzn-3	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.7.1-amzn-5	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.11.1	A biblioteca de Machine Learning.
mysql-server	5.5	O servidor de banco de dados MySQL.

Componente	Versão	Descrição
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
presto-coordinator	0.136	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.136	O serviço que executa partes de uma consulta.
pig-client	0.14.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	1.6.0	Os clientes da linha de comando do Spark.
spark-history-server	1.6.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	1.6.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	1.6.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
sqoop-client	1.4.6	O cliente da linha de comando do Apache Sqoop.

Componente	Versão	Descrição
webserver	2.4	O servidor Apache HTTP.
zeppelin-server	0.5.6-incubating	O notebook baseado na web que permite um data analytics interativo.

Classificações de configuração 4.4.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-4.4.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hcatalog-env	Mude os valores no HCatalog ambiente.
hcatalog-server-jndi	Altere os valores em HCatalog jndi.properties.

Classificações	Descrição
hcatalog-server-proto-hive-site	Altere os valores em HCatalog proto-hive-site .xml.
hcatalog-webhcat-env	Mude os valores no ambiente HCat da HCatalog Web.
hcatalog-webhcat-log4j	Altere os valores nas HCatalog propriedades HCat log4j da Web.
hcatalog-webhcat-site	Altera os valores no arquivo webhcat-site.xml HCat do HCatalog site da Web.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j	Altere os valores no arquivo hive-exec-log 4j.properties do Hive.
hive-log4j	Alterar os valores no arquivo hive-log4j.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.

Classificações	Descrição
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altera os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.

Classificações	Descrição
spark-env	Alterar os valores no ambiente do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
sqoop-env	Alterar os valores no ambiente do Sqoop.
sqoop-oraoop-site	Altera os valores no arquivo oraoop-site.xml OraOop do Sqoop.
sqoop-site	Alterar os valores no arquivo sqoop-site.xml do Sqoop.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.

Versão 4.3.0 do Amazon EMR

Versões 4.3.0 da aplicação

Esta versão inclui os seguintes aplicativos: GangliaHadoop, Hive, Hue, Mahout, Oozie-SandboxPig, Presto-Sandbox, Spark,, Zeppelin-Sandboxe.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR

- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
AWS SDK for Java	1.10.27	1.10.27	Não rastreadas	Não rastreadas
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.6.0	-	-
HBase	-	-	-	-
HCatalog	-	-	-	-
Hadoop	2.7.1	2.6.0	2.6.0	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	-
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.0	0.11.0	0.11.0	0.10.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.0.1	-
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.130	0.125	0.119	-
Spark	1.6.0	1.5.2	1.5.0	1.4.1
Sqoop	-	-	-	-
Sqoop-Sandbox	-	-	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.5	0.5.5	0.6.0-SNAPSHOT	-
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

Notas da versão 4.3.0

As notas da versão a seguir incluem informações para a versão 4.3.0 do Amazon EMR.

Data do release: 19 de janeiro de 2016

Atributos

- Atualizado para Hadoop 2.7.1
- Atualizado para Spark 1.6.0
- Ganglia atualizado para 3.7.2
- Presto atualizado para 0.130
- O Amazon EMR fez algumas alterações na configuração `spark.dynamicAllocation.enabled` quando ela estava definida como verdadeiro. Ela é definida como falso por padrão. Quando definida como verdadeiro, isso afeta os padrões definidos pela configuração `maximizeResourceAllocation`:
  - Se `spark.dynamicAllocation.enabled` estiver definida como true, `spark.executor.instances` não será definida por `maximizeResourceAllocation`.
  - A configuração `spark.driver.memory` agora é definida com base nos tipos de instância no cluster de maneira semelhante a como `spark.executors.memory` é definida. No entanto, como o aplicativo de driver do Spark pode ser executado na instância principal ou em uma das instâncias core (por exemplo, nos modos de cluster e de cliente do YARN, respectivamente), a configuração `spark.driver.memory` é definida com base no tipo de instância do menor tipo de instância entre esses dois grupos de instâncias.
  - A configuração `spark.default.parallelism` agora é definida em duas vezes mais núcleos de CPU disponíveis para contêineres do YARN. Em versões anteriores, era a metade desse valor.
  - Os cálculos para a sobrecarga de memória reservada para os processos do Spark YARN foram ajustados para serem mais precisos, resultando em um pequeno aumento na quantidade total de memória disponível para o Spark (ou seja, `spark.executor.memory`).

Problemas conhecidos das versões anteriores que foram resolvidos

- A agregação de logs do YARN agora está ativada por padrão.
- Corrigido um problema em que os logs não eram enviados para um bucket de logs do Amazon S3 do cluster quando a agregação de logs do YARN estava ativada.

- Os contêineres do YARN agora têm um novo mínimo de 32 em todos os tipos de nós.
- Corrigido um problema com o Ganglia que causava E/S de disco excessivas no nó primário em clusters grandes.
- Corrigido um problema que impedia que os logs de aplicações fossem enviados para o Amazon S3 quando um cluster estivesse sendo encerrado.
- Corrigido um problema na CLI do EMRFS que fazia com que houvesse falha em determinados comandos.
- Corrigido um problema com o Zeppelin que impedia que dependências fossem carregadas no subjacente. SparkContext
- Corrigido um problema resultante da emissão de um redimensionamento para tentar adicionar instâncias.
- Corrigido um problema no Hive, em que CREATE TABLE AS SELECT fazia chamadas de lista excessivas para o Amazon S3.
- Corrigido um problema em que clusters grandes não provisionavam corretamente quando o Hue, o Oozie e o Ganglia estivessem instalados.
- Corrigido um problema no s3-dist-cp em que retornava um código de saída zero, mesmo se houvesse falha com um erro.

Versões de componente 4.3.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	3.0.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.0.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.1.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.1.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.3.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
ganglia-monitor	3.7.2	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.7.2	O coletor de metadados do Ganglia para agregar métricas de agentes de monitoramento do Ganglia.
ganglia-web	3.7.1	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.

Componente	Versão	Descrição
hadoop-client	2.7.1-amzn-0	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.7.1-amzn-0	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.7.1-amzn-0	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.7.1-amzn-0	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.7.1-amzn-0	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.7.1-amzn-0	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er
hadoop-mapred	2.7.1-amzn-0	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.7.1-amzn-0	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.7.1-amzn-0	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.

Componente	Versão	Descrição
hive-client	1.0.0-amzn-2	O cliente da linha de comando do Hive.
hive-metastore-server	1.0.0-amzn-2	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	1.0.0-amzn-2	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.7.1-amzn-5	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.11.0	A biblioteca de Machine Learning.
mysql-server	5.5	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.
oozie-server	4.2.0	O serviço que aceita solicitaç ões de fluxo de trabalho do Oozie.
presto-coordinator	0.130	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-wo rkers.
presto-worker	0.130	O serviço que executa partes de uma consulta.

Componente	Versão	Descrição
pig-client	0.14.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	1.6.0	Os clientes da linha de comando do Spark.
spark-history-server	1.6.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	1.6.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	1.6.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
webserver	2.4	O servidor Apache HTTP.
zeppelin-server	0.5.5-incubating-amzn-1	O notebook baseado na web que permite um data analytics interativo.

Classificações de configuração 4.3.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-4.3.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.

Classificações	Descrição
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j	Altere os valores no arquivo hive-exec-log 4j.properties do Hive.
hive-log4j	Alterar os valores no arquivo hive-log4j.propert ies do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hue-ini	Alterar os valores no arquivo ini do Hue
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.

Classificações	Descrição
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altera os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.

Classificações	Descrição
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.

Versão 4.2.0 do Amazon EMR

Versões 4.2.0 da aplicação

Esta versão inclui os seguintes aplicativos: GangliaHadoop, Hive, Hue, Mahout, Oozie-SandboxPig, Presto-Sandbox, Spark,, Zeppelin-Sandboxe.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR

- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
AWS SDK for Java	1.10.27	1.10.27	Não rastreadas	Não rastreadas
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.6.0	-	-
HBase	-	-	-	-
HCatalog	-	-	-	-
Hadoop	2.7.1	2.6.0	2.6.0	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	-
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.0	0.11.0	0.11.0	0.10.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.0.1	-
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.130	0.125	0.119	-
Spark	1.6.0	1.5.2	1.5.0	1.4.1
Sqoop	-	-	-	-
Sqoop-Sandbox	-	-	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.5	0.5.5	0.6.0-SNAPSHOT	-
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

Notas da versão 4.2.0

As notas da versão a seguir incluem informações para a versão 4.2.0 do Amazon EMR.

Data do release: 18 de novembro de 2015

Atributos

- Adicionado o suporte ao Ganglia
- Atualizado para Spark 1.5.2
- Atualizado para Presto 0.125
- Oozie atualizado para 4.2.0
- Zeppelin atualizado para 0.5.5
- Atualizou o AWS SDK para Java para 1.10.27

Problemas conhecidos das versões anteriores que foram resolvidos

- Corrigido um problema com a CLI do EMRFS em que ela não usava o nome da tabela de metadados padrão.
- Corrigido um problema encontrado ao serem usadas tabelas com ORC no Amazon S3.
- Corrigido um problema encontrado com uma divergência de versão do Python na configuração do Spark.
- Corrigido um problema quando havia falha de um status de nó YARN em relatar problemas de DNS para clusters em uma VPC.
- Corrigido um problema encontrado quando o YARN desativava nós, resultando em aplicações suspensas ou na incapacidade de agendar novas aplicações.
- Corrigido um problema encontrado quando os clusters eram encerrados com o status TIMED_OUT_STARTING.
- Corrigido um problema encontrado quando dependência Scala do EMRFS Scala era incluída em outras compilações. A dependência Scala foi removida.

Versões de componente 4.2.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`.

Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *CommunityVersion*-amzn-*EmrVersion*. O *EmrVersion* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	3.0.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.0.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.1.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.0.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.2.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
ganglia-monitor	3.6.0	O agente incorporado do Ganglia para aplicativos do ecossistema do Hadoop, juntamente com o agente de monitoramento do Ganglia.
ganglia-metadata-collector	3.6.0	O coletor de metadados do Ganglia para agregar métricas

Componente	Versão	Descrição
		de agentes de monitoramento do Ganglia.
ganglia-web	3.5.10	O aplicativo web para visualizar as métricas coletadas pelo coletor de metadados do Ganglia.
hadoop-client	2.6.0-amzn-2	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.6.0-amzn-2	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.6.0-amzn-2	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.6.0-amzn-2	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.6.0-amzn-2	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.6.0-amzn-2	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvider
hadoop-mapred	2.6.0-amzn-2	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.

Componente	Versão	Descrição
hadoop-yarn-nodemanager	2.6.0-amzn-2	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.6.0-amzn-2	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hive-client	1.0.0-amzn-1	O cliente da linha de comando do Hive.
hive-metastore-server	1.0.0-amzn-1	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	1.0.0-amzn-1	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.7.1-amzn-5	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.11.0	A biblioteca de Machine Learning.
mysql-server	5.5	O servidor de banco de dados MySQL.
oozie-client	4.2.0	O cliente da linha de comando do Oozie.

Componente	Versão	Descrição
oozie-server	4.2.0	O serviço que aceita solicitações de fluxo de trabalho do Oozie.
presto-coordinator	0.125	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.125	O serviço que executa partes de uma consulta.
pig-client	0.14.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	1.5.2	Os clientes da linha de comando do Spark.
spark-history-server	1.5.2	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	1.5.2	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	1.5.2	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
webserver	2.4	O servidor Apache HTTP.
zeppelin-server	0.5.5-incubating-amzn-0	O notebook baseado na web que permite um data analytics interativo.

Classificações de configuração 4.2.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-4.2.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j	Altere os valores no arquivo hive-exec-log4j.properties do Hive.
hive-log4j	Alterar os valores no arquivo hive-log4j.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hue-ini	Alterar os valores no arquivo ini do Hue

Classificações	Descrição
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altera os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.

Classificações	Descrição
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
presto-connector-hive	Alterar os valores no arquivo hive.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
spark-metrics	Alterar os valores no arquivo metrics.properties do Spark.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.

Versão 4.1.0 do Amazon EMR

Versões 4.1.0 da aplicação

Esta versão inclui os seguintes aplicativos: HadoopHive, Hue, Mahout, Oozie-SandboxPig, Presto-Sandbox, Spark,, Zeppelin-Sandboxe.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR
- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
AWS SDK for Java	1.10.27	1.10.27	Não rastreadas	Não rastreadas
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.6.0	-	-
HBase	-	-	-	-
HCatalog	-	-	-	-
Hadoop	2.7.1	2.6.0	2.6.0	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
Hue	3.7.1	3.7.1	3.7.1	-
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.0	0.11.0	0.11.0	0.10.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.0.1	-
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.130	0.125	0.119	-
Spark	1.6.0	1.5.2	1.5.0	1.4.1
Sqoop	-	-	-	-
Sqoop-Sandbox	-	-	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.5	0.5.5	0.6.0-SNAPSHOT	-
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

Notas da versão 4.1.0

Versões de componente 4.1.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	3.0.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.0.0	Bibliotecas convenientes para o ecossistema do Hadoop.

Componente	Versão	Descrição
emr-kinesis	3.1.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.0.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.1.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.
hadoop-client	2.6.0-amzn-1	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.6.0-amzn-1	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-library	2.6.0-amzn-1	O cliente de linha de comando e biblioteca do HDFS
hadoop-hdfs-namenode	2.6.0-amzn-1	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.6.0-amzn-1	O endpoint de HTTP para as operações do HDFS.
hadoop-kms-server	2.6.0-amzn-1	Servidor de gerenciamento de chaves criptográficas baseado na API do Hadoop. KeyProvid er

Componente	Versão	Descrição
hadoop-mapred	2.6.0-amzn-1	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.6.0-amzn-1	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.6.0-amzn-1	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hive-client	1.0.0-amzn-1	O cliente da linha de comando do Hive.
hive-metastore-server	1.0.0-amzn-1	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.
hive-server	1.0.0-amzn-1	O serviço que aceita as consultas do Hive como solicitações da web.
hue-server	3.7.1-amzn-4	O aplicativo web para analisar dados usando aplicativos do ecossistema do Hadoop
mahout-client	0.11.0	A biblioteca de Machine Learning.
mysql-server	5.5	O servidor de banco de dados MySQL.

Componente	Versão	Descrição
oozie-client	4.0.1	O cliente da linha de comando do Oozie.
oozie-server	4.0.1	O serviço que aceita solicitações de fluxo de trabalho do Oozie.
presto-coordinator	0.119	O serviço que aceita consultas e gerencia a execução de consultas entre os presto-workers.
presto-worker	0.119	O serviço que executa partes de uma consulta.
pig-client	0.14.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	1.5.0	Os clientes da linha de comando do Spark.
spark-history-server	1.5.0	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	1.5.0	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	1.5.0	As bibliotecas do Apache Spark necessárias para subordinados do YARN.
zeppelin-server	0.6.0-incubating-SNAPSHOT	O notebook baseado na web que permite um data analytics interativo.

Classificações de configuração 4.1.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-4.1.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hdfs-encryption-zones	Configurar as zonas de criptografia do HDFS.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j	Altere os valores no arquivo hive-exec-log4j.properties do Hive.
hive-log4j	Alterar os valores no arquivo hive-log4j.properties do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
hue-ini	Alterar os valores no arquivo ini do Hue

Classificações	Descrição
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
hadoop-kms-acls	Alterar os valores no arquivo kms-acls.xml do Hadoop.
hadoop-kms-env	Alterar os valores no ambiente do Hadoop KMS.
hadoop-kms-log4j	Alterar os valores no arquivo kms-log4j.properties do Hadoop.
hadoop-kms-site	Alterar os valores no arquivo kms-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.
mapred-site	Altera os valores no arquivo mapred-site.xml do MapReduce aplicativo.
oozie-env	Alterar os valores no ambiente do Oozie.
oozie-log4j	Alterar os valores no arquivo oozie-log4j.properties do Oozie.
oozie-site	Alterar os valores no arquivo oozie-site.xml do Oozie.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.

Classificações	Descrição
presto-log	Alterar os valores no arquivo log.properties do Presto.
presto-config	Alterar os valores no arquivo config.properties do Presto.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.
zeppelin-env	Alterar os valores no ambiente do Zeppelin.

Versão 4.0.0 do Amazon EMR

Versões 4.0.0 da aplicação

Esta versão inclui os seguintes aplicativos: HadoopHiveMahout, Pig,, Sparke.

A tabela abaixo lista as versões das aplicações disponíveis nesta versão do Amazon EMR e as versões de aplicações nas três versões anteriores do Amazon EMR (quando aplicável).

Para obter um histórico abrangente das versões das aplicações de cada versão do Amazon EMR, consulte os seguintes tópicos:

- Versões de aplicações nas versões 7.x do Amazon EMR
- Versões de aplicações nas versões 6.x do Amazon EMR

- Versões de aplicações nas versões 5.x do Amazon EMR
- Versões de aplicações nas versões 4.x do Amazon EMR

Informações da versão da aplicação

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
AWS SDK for Java	1.10.27	1.10.27	Não rastreadas	Não rastreadas
Python	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
Scala	Não rastreadas	Não rastreadas	Não rastreadas	Não rastreadas
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.6.0	-	-
HBase	-	-	-	-
HCatalog	-	-	-	-
Hadoop	2.7.1	2.6.0	2.6.0	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	-
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.0	0.11.0	0.11.0	0.10.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.0.1	-
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.130	0.125	0.119	-
Spark	1.6.0	1.5.2	1.5.0	1.4.1
Sqoop	-	-	-	-
Sqoop-Sandbox	-	-	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.5	0.5.5	0.6.0-SNAPSHOT	-
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

Notas da versão 4.0.0

Versões de componente 4.0.0

Os componentes que o Amazon EMR instala com esta versão estão listados abaixo. Alguns são instalados como parte de pacotes de aplicativos de big data. Outros são exclusivos do Amazon EMR e instalados para processos e atributos do sistema. Eles normalmente começam com `emr` ou `aws`. Os pacotes de aplicações de big data na versão mais recente do Amazon EMR são geralmente a versão mais recente encontrada na comunidade. Disponibilizamos as versões da comunidade no Amazon EMR o mais rapidamente possível.

Alguns componentes no Amazon EMR diferem das versões da comunidade. Esses componentes tem um rótulo de versão no formulário *`CommunityVersion`*`-amzn-`*`EmrVersion`*. O *`EmrVersion`* começa em 0. Por exemplo, se um componente da comunidade de código aberto denominado `myapp-component` com a versão 2.2 tiver sido alterado três vezes para inclusão em versões diferentes do Amazon EMR, sua versão será listada como `2.2-amzn-2`.

Componente	Versão	Descrição
emr-ddb	3.0.0	O conector do Amazon DynamoDB para aplicativos do ecossistema do Hadoop.
emr-goodies	2.0.0	Bibliotecas convenientes para o ecossistema do Hadoop.
emr-kinesis	3.0.0	O conector do Amazon Kinesis para aplicativos do ecossistema do Hadoop.
emr-s3-dist-cp	2.0.0	Cópia distribuída otimizada de aplicativos para o Amazon S3.
emrfs	2.0.0	O conector do Amazon S3 para aplicações do ecossiste ma do Hadoop.

Componente	Versão	Descrição
hadoop-client	2.6.0-amzn-0	Clientes da linha de comando do Hadoop, como 'hdfs', 'hadoop', ou 'yarn'.
hadoop-hdfs-datanode	2.6.0-amzn-0	O serviço de nível de nó do HDFS para armazenamento de blocos.
hadoop-hdfs-namenode	2.6.0-amzn-0	O serviço do HDFS para rastrear nomes de arquivos e locais de blocos.
hadoop-httpfs-server	2.6.0-amzn-0	O endpoint de HTTP para as operações do HDFS.
hadoop-mapred	2.6.0-amzn-0	MapReduce bibliotecas de mecanismos de execução para executar um MapReduce aplicativo.
hadoop-yarn-nodemanager	2.6.0-amzn-0	O serviço do YARN para o gerenciamento de contêineres em um nó individual.
hadoop-yarn-resour cemanager	2.6.0-amzn-0	O serviço do YARN para alocar e gerenciar recursos de cluster e aplicativos distribuí dos.
hive-client	1.0.0-amzn-0	O cliente da linha de comando do Hive.
hive-metastore-server	1.0.0-amzn-0	O serviço de acesso à metastore do Hive, um repositório de semântica que armazena metadados do SQL nas operações do Hadoop.

Componente	Versão	Descrição
hive-server	1.0.0-amzn-0	O serviço que aceita as consultas do Hive como solicitações da web.
mahout-client	0.10.0	A biblioteca de Machine Learning.
mysql-server	5.5	O servidor de banco de dados MySQL.
pig-client	0.14.0-amzn-0	O cliente da linha de comando do Pig.
spark-client	1.4.1	Os clientes da linha de comando do Spark.
spark-history-server	1.4.1	A interface de usuário da web para visualizar os eventos registrados por toda a vida útil de um aplicativo Spark concluído.
spark-on-yarn	1.4.1	O mecanismo de execução em memória para o YARN.
spark-yarn-slave	1.4.1	As bibliotecas do Apache Spark necessárias para subordinados do YARN.

Classificações de configuração 4.0.0

As classificações de configuração permitem que você personalize aplicações. Elas geralmente correspondem a um arquivo XML de configuração da aplicação, como `hive-site.xml`. Para obter mais informações, consulte Configurar aplicações.

Classificações do emr-4.0.0

Classificações	Descrição
capacity-scheduler	Alterar os valores no arquivo capacity-scheduler.xml do Hadoop.
core-site	Alterar os valores no arquivo core-site.xml do Hadoop.
emrfs-site	Alterar as configurações do EMRFS.
hadoop-env	Alterar os valores no ambiente do Hadoop para todos os componentes do Hadoop.
hadoop-log4j	Alterar os valores no arquivo log4j.properties do Hadoop.
hdfs-site	Alterar os valores no arquivo hdfs-site.xml do HDFS.
hive-env	Alterar os valores no ambiente do Hive.
hive-exec-log4j	Altere os valores no arquivo hive-exec-log 4j.properties do Hive.
hive-log4j	Alterar os valores no arquivo hive-log4j.propert ies do Hive.
hive-site	Alterar os valores no arquivo hive-site.xml do Hive.
httpfs-env	Alterar os valores no ambiente do HTTPFS.
httpfs-site	Alterar os valores no arquivo httpfs-site.xml do Hadoop.
mapred-env	Altere os valores no ambiente do MapReduce aplicativo.

Classificações	Descrição
mapred-site	Altera os valores no arquivo mapred-site.xml do MapReduce aplicativo.
pig-properties	Alterar os valores no arquivo pig.properties do Pig.
pig-log4j	Alterar os valores no arquivo log4j.properties do Pig.
spark	As configurações administradas pelo Amazon EMR para o Apache Spark.
spark-defaults	Alterar os valores no arquivo spark-defaults.conf do Spark.
spark-env	Alterar os valores no ambiente do Spark.
spark-log4j	Alterar os valores no arquivo log4j.properties do Spark.
yarn-env	Alterar os valores no ambiente do YARN.
yarn-site	Alterar os valores no arquivo yarn-site.xml do YARN.

Versões 2.x e 3.x da AMI do Amazon EMR

> **Note**
>
> AWS A está atualizando a configuração de TLS para todos os endpoints AWS da API da para uma versão mínima do TLS 1.2. As versões 3.10 e anteriores do Amazon EMR são compatíveis somente com conexões TLS 1.0/1.1. Depois de 4 de dezembro de 2023, você não poderá criar clusters com as versões 3.10 e anteriores do Amazon EMR.

> Se você usa as versões 3.10 ou anteriores do Amazon EMR, recomendamos que você teste e migre imediatamente suas workloads para a versão mais recente do Amazon EMR. Para obter mais informações, consulte o Blog de segurança da AWS.

As versões 2.x e 3.x do Amazon EMR, denominadas versões da AMI, são disponibilizadas para soluções preexistentes que as exigem por motivos de compatibilidade. Não recomendamos a criação de novos clusters ou novas soluções com essas versões. Elas não têm recursos de versões mais recentes e incluem pacotes de aplicativos desatualizados.

Recomendamos que você crie soluções usando a versão mais recente do Amazon EMR.

O grau de diferença entre as versões das séries 2.x e 3.x e as versões recentes do Amazon EMR é significativo. Essas diferenças vão desde como criar e configurar um cluster para as portas e a estrutura do diretório dos aplicativos no cluster.

Esta seção tenta abordar as diferenças mais significativas do Amazon EMR, bem como as diferenças específicas de configuração e gerenciamento de aplicações. Ela não é abrangente. Se você criar e usar clusters nas séries 2.x ou 3.x, poderá encontrar diferenças não abordadas nesta seção.

Tópicos

- Criar um cluster com versões anteriores da AMI do Amazon EMR
- Instalar aplicações com versões anteriores da AMI do Amazon EMR
- Personalizar clusters e configurações de aplicações com versões anteriores da AMI do Amazon EMR
- Detalhes da aplicação Hive para versões anteriores da AMI do Amazon EMR
- HBase detalhes da aplicação para versões anteriores da AMI do Amazon EMR
- Detalhes da aplicação Pig para versões anteriores da AMI do Amazon EMR
- Detalhes de aplicações do Spark com versões anteriores da AMI do Amazon EMR
- Diferenças do DistCp utilitário S3 com versões anteriores da AMI do Amazon EMR

Criar um cluster com versões anteriores da AMI do Amazon EMR

As versões 2.x e 3.x do Amazon EMR são conhecidas como versão da AMI. Desde as versões 4.0.0 e posteriores do Amazon EMR, as versões são designadas por um rótulo de versão, como

`emr-5.11.0`. Essa alteração é mais aparente quando você cria um cluster usando o AWS CLI ou programaticamente.

Ao usar o AWS CLI para criar um cluster usando uma versão de lançamento da AMI, use a `--ami-version` opção, por exemplo,`--ami-version 3.11.0`. Muitas opções, atributos e aplicações apresentadas nas versões 4.0.0 e posteriores do Amazon EMR não estão disponíveis quando você especifica uma `--ami-version`. Para obter mais informações, consulte create-cluster na AWS CLI Command Reference.

O AWS CLI comando de exemplo a seguir inicia um cluster usando uma versão da AMI.

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr create-cluster --name "Test cluster" --ami-version 3.11.0 \
--applications Name=Hue Name=Hive Name=Pig \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-groups InstanceGroupType=MASTER,InstanceCount=1,\
InstanceType=m3.xlarge InstanceGroupType=CORE,InstanceCount=2,\
InstanceType=m3.xlarge --bootstrap-actions Path=s3://elasticmapreduce/bootstrap-
actions/configure-hadoop,\
Name="Configuring infinite JVM reuse",Args=["-m","mapred.job.reuse.jvm.num.tasks=-1"]
```

Programaticamente, todas as versões do Amazon EMR usam a ação `RunJobFlowRequest` na API do EMR para criar clusters. O seguinte exemplo de código Java cria um cluster usando a versão da AMI 3.11.0.

```
RunJobFlowRequest request = new RunJobFlowRequest()
   .withName("AmiVersion Cluster")
   .withAmiVersion("3.11.0")
   .withInstances(new JobFlowInstancesConfig()
    .withEc2KeyName("myKeyPair")
    .withInstanceCount(1)
    .withKeepJobFlowAliveWhenNoSteps(true)
    .withMasterInstanceType("m3.xlarge")
    .withSlaveInstanceType("m3.xlarge");
```

A chamada `RunJobFlowRequest` usa um rótulo de versão:

```
RunJobFlowRequest request = new RunJobFlowRequest()
   .withName("ReleaseLabel Cluster")
   .withReleaseLabel("emr-7.9.0")
   .withInstances(new JobFlowInstancesConfig()
    .withEc2KeyName("myKeyPair")
    .withInstanceCount(1)
    .withKeepJobFlowAliveWhenNoSteps(true)
    .withMasterInstanceType("m3.xlarge")
    .withSlaveInstanceType("m3.xlarge");
```

Configurar o tamanho do cluster

Durante a execução do cluster, o Hadoop determina o número de tarefas de mapeador e reducer necessárias para processar os dados. Clusters maiores devem ter mais tarefas para fazer melhor utilização dos recursos e reduzir o tempo de processamento. Normalmente, um cluster do EMR permanece do mesmo tamanho durante toda a execução do cluster; você define o número de tarefas ao criar o cluster. Quando você redimensiona um cluster em execução, pode variar o processamento durante a execução do cluster. Portanto, em vez de usar um número fixo de tarefas, você pode variar o número de tarefas durante a vida útil do cluster. Há duas opções de configuração para ajudar a definir o número ideal de tarefas:

- `mapred.map.tasksperslot`
- `mapred.reduce.tasksperslot`

Você pode definir ambas as opções no arquivo `mapred-conf.xml`. Quando você envia um trabalho para o cluster, o cliente do trabalho verifica o número total de slots de mapeamento e redução em todo o cluster. O cliente do trabalho, em seguida, usa as seguintes equações para definir o número de tarefas:

- `mapred.map.tasks` = `mapred.map.tasksperslot` * slots de mapeamento no cluster
- `mapred.reduce.tasks` = `mapred.reduce.tasksperslot` * slots de redução no cluster

O cliente do trabalho lê o parâmetro `tasksperslot` somente se o número de tarefas não estiver configurado. Você pode substituir o número de tarefas a qualquer momento para todos os clusters por meio de uma ação de bootstrap, ou pode fazê-lo individualmente por trabalho, adicionando uma etapa para alterar a configuração.

O Amazon EMR é capaz de superar falhas de nós de tarefas e prosseguir com a execução do cluster, mesmo que o nó de tarefa se torne indisponível. O Amazon EMR disponibiliza automaticamente nós de tarefa adicionais para substituir aqueles que falham.

Você pode ter um número diferente de nós de tarefa para cada etapa do cluster. Você também pode adicionar uma etapa em um cluster em execução para modificar o número de nós de tarefa. Como é certo que todas as etapas são executadas sequencialmente por padrão, você pode especificar o número de nós de tarefa em execução para qualquer etapa.

Instalar aplicações com versões anteriores da AMI do Amazon EMR

Ao usar uma versão da AMI, os aplicativos são instalados de várias maneiras, inclusive usando o `NewSupportedProducts` parâmetro para a RunJobFlowação, usando ações de bootstrap e usando a ação Step.

Personalizar clusters e configurações de aplicações com versões anteriores da AMI do Amazon EMR

A versão 4.0.0 do Amazon EMR introduziu um método simplificado de configuração de aplicações usando classificações de configuração. Para obter mais informações, consulte Configurar aplicações. Ao usar uma versão da AMI, você configura aplicativos usando ações de bootstrap juntamente com os argumentos que transmite. Por exemplo, as ações de bootstrap `configure-hadoop` e `configure-daemons` definem propriedades do ambiente específicas do Hadoop e do YARN, como `--namenode-heap-size`. Em versões mais recentes, elas são configuradas usando as classificações de configuração `hadoop-env` e `yarn-env`. Para ações de bootstrap que realizam configurações comuns, consulte o emr-bootstrap-actions repositório no Github.

As tabelas a seguir mapeiam ações de bootstrap para classificações de configuração em versões mais recentes do Amazon EMR.

Hadoop

Nome do arquivo de aplicativo afetado	Ação de bootstrap da versão da AMI	Classificação de configuração
`core-site.xml`	`configure-hadoop -c`	`core-site`
`log4j.properties`	`configure-hadoop -l`	`hadoop-log4j`

Nome do arquivo de aplicativo afetado	Ação de bootstrap da versão da AMI	Classificação de configuração
hdfs-site.xml	configure-hadoop -s	hdfs-site
n/a	n/a	hdfs-encryption-zones
mapred-site.xml	configure-hadoop -m	mapred-site
yarn-site.xml	configure-hadoop -y	yarn-site
httpfs-site.xml	configure-hadoop -t	httpfs-site
capacity-scheduler.xml	configure-hadoop -z	capacity-scheduler
yarn-env.sh	configure-daemons --resourcemanager-opts	yarn-env

Hive

Nome do arquivo de aplicativo afetado	Ação de bootstrap da versão da AMI	Classificação de configuração
hive-env.sh	n/a	hive-env
hive-site.xml	hive-script --install-hive-site ${MY_HIVE_SITE_FILE}	hive-site
hive-exec-log4j.properties	n/a	hive-exec-log4j
hive-log4j.properties	n/a	hive-log4j

EMRFS

Nome do arquivo de aplicativo afetado	Ação de bootstrap da versão da AMI	Classificação de configuração
`emrfs-site.xml`	`configure-hadoop -e`	`emrfs-site`
n/a	`s3get -s s3://custom-provider.jar -d /usr/share/aws/emr/auxlib/`	`emrfs-site` (com nova configuração `fs.s3.cse.encryptionMaterialsProvider.uri`)

Para obter uma lista de todas as classificações, consulte Configurar aplicações.

Variáveis de ambientes de aplicações

Quando você usa uma versão da AMI, um script `hadoop-user-env.sh` é usado juntamente com a ação de bootstrap `configure-daemons` para configurar o ambiente do Hadoop. O script inclui as seguintes ações:

```
#!/bin/bash
export HADOOP_USER_CLASSPATH_FIRST=true;
echo "HADOOP_CLASSPATH=/path/to/my.jar" >> /home/hadoop/conf/hadoop-user-env.sh
```

Na versão 4.x do Amazon EMR, você faz o mesmo usando a classificação de configuração `hadoop-env`, conforme mostrado no exemplo a seguir:

```
[
      {
         "Classification":"hadoop-env",
         "Properties":{

         },
         "Configurations":[
            {
               "Classification":"export",
               "Properties":{
                  "HADOOP_USER_CLASSPATH_FIRST":"true",
                  "HADOOP_CLASSPATH":"/path/to/my.jar"
               }
```

```
            }
         ]
      }
   ]
```

Como outro exemplo, usar `configure-daemons` e transmitir `--namenode-heap-size=2048` e `--namenode-opts=-XX:GCTimeRatio=19` é equivalente às seguintes classificações de configuração.

```
[
      {
         "Classification":"hadoop-env",
         "Properties":{

         },
         "Configurations":[
            {
               "Classification":"export",
               "Properties":{
                  "HADOOP_DATANODE_HEAPSIZE":  "2048",
            "HADOOP_NAMENODE_OPTS":  "-XX:GCTimeRatio=19"
               }
            }
         ]
      }
   ]
```

As outras variáveis do ambiente de aplicativos não são mais definidas em `/home/hadoop/.bashrc`. Em vez disso, elas são definidas nos arquivos `/etc/default` por componente ou aplicativo, como por exemplo `/etc/default/hadoop`. Os scripts de wrapper `/usr/bin/` instalados pelo aplicativo também RPMs podem definir variáveis de ambiente adicionais antes de envolver o script bin real.

Portas de serviço

Ao usar uma versão da AMI, alguns serviços usam portas personalizadas.

Alterações nas configurações das portas

Configuração	AMI versão 3.x	Padrão de código aberto
fs.default.name	hdfs://emrDeterminedIP:9000	padrão (hdfs: *emrDeterminedIP* //:8020)
dfs.datanode.address	0.0.0.0:9200	default (0.0.0.0:50010)
dfs.datanode.http.address	0.0.0.0:9102	default (0.0.0.0:50075)
dfs.datanode.https.address	0.0.0.0:9402	default (0.0.0.0:50475)
dfs.datanode.ipc.address	0.0.0.0:9201	default (0.0.0.0:50020)
dfs.http.address	0.0.0.0:9101	default (0.0.0.0:50070)
dfs.https.address	0.0.0.0:9202	default (0.0.0.0:50470)
dfs.secondary.http.address	0.0.0.0:9104	default (0.0.0.0:50090)
yarn.nodemanager.address	0.0.0.0:9103	default (${yarn.nodemanager.hostname}:0)
yarn.nodemanager.localizer.address	0.0.0.0:9033	default (${yarn.nodemanager.hostname}:8040)
yarn.nodemanager.webapp.address	0.0.0.0:9035	default (${yarn.nodemanager.hostname}:8042)
yarn.resourcemanager.address	*emrDeterminedIP* : 9022	default (${yarn.resourcemanager.hostname}:8032)
yarn.resourcemanager.admin.address	*emrDeterminedIP* : 9025	default (${yarn.resourcemanager.hostname}:8033)
yarn.resourcemanager.resource-tracker.address	*emrDeterminedIP* : 9023	default (${yarn.resourcemanager.hostname}:8031)
yarn.resourcemanager.scheduler.address	*emrDeterminedIP* : 9024	default (${yarn.resourcemanager.hostname}:8030)

Configuração	AMI versão 3.x	Padrão de código aberto
yarn.resourcemanager.webapp.address	0.0.0.0:9026	default (${yarn.resourcemanager.hostname}:8088)
yarn.web-proxy.address	*emrDeterminedIP* : 9046	default (no-value)
yarn.resourcemanager.hostname	0.0.0.0 (default)	*emrDeterminedIP*

> **Note**
>
> *emrDeterminedIP*É um endereço IP gerado pelo Amazon EMR.

Usuários

Ao usar uma versão da AMI, o usuário hadoop executa todos os processos e é proprietário de todos os arquivos. Nas versões 4.0.0 e posteriores do Amazon EMR, existem usuários no nível da aplicação e dos componentes.

Sequência de instalação, artefatos instalados e locais dos arquivos de log

Ao usar uma versão da AMI, os artefatos de aplicativos e seus diretórios de configuração são instalados no diretório /home/hadoop/*application*. Por exemplo, se você tivesse o Hive instalado, o diretório seria /home/hadoop/hive. Nas versões 4.0.0 e posteriores do Amazon EMR, os artefatos de aplicações são instalados no diretório /usr/lib/*application*. Ao usar uma versão da AMI, os arquivos de log são encontrados em vários locais. A tabela a seguir lista os locais.

Alterações nos locais de log no Amazon S3

Daemon ou aplicação	Local do diretório
instance-state	nodo/ /instance-state/ *instance-id*
hadoop-hdfs-namenode	*instance-id* daemons///.log hadoop-hadoop-namenode

Daemon ou aplicação	Local do diretório
hadoop-hdfs-datanode	*instance-id* daemons///.log hadoop-hadoop-datanode
hadoop-yarn () ResourceManager	*instance-id* daemons//yarn-hadoop-resourcemanager
hadoop-yarn (Proxy Server)	*instance-id* daemons//yarn-hadoop-proxyserver
mapred-historyserver	*instance-id* daemons//
httpfs	*instance-id* daemons/ /httpfs.log
hive-server	node/ /hive-server/hive-server.log *instance-id*
hive-metastore	node/ /apps/hive.log *instance-id*
CLI do Hive	node/ /apps/hive.log *instance-id*
Logs de usuário e logs de contêiner dos aplicativos YARN	task-attempts/
Mahout	N/D
Pig	N/D
spark-historyserver	N/D
Arquivos de histórico de trabalhos do mapreduce	jobs/

Executor de comandos

Quando uma versão da AMI é usada, muitos scripts ou programas como, por exemplo `/home/hadoop/contrib/streaming/hadoop-streaming.jar`, não são colocados no ambiente do caminho de login do shell. Portanto, você precisa especificar o caminho completo ao usar um arquivo jar, como command-runner.jar ou script-runner.jar, para executar os scripts. O `command-`

`runner.jar` está localizado na AMI; portanto, não há necessidade de conhecer um URI completo, como era o caso com `script-runner.jar`.

Fator de replicação

O fator de replicação permite que você configure quando uma JVM do Hadoop deverá ser iniciada. Você pode iniciar uma nova JVM do Hadoop para cada tarefa, o que fornece melhor isolamento de tarefas, ou você pode compartilhar JVMs entre tarefas, fornecendo menor sobrecarga da estrutura. Se você estiver processando vários arquivos pequenos, faz sentido reutilizar o JVM muitas vezes para amortizar o custo de startup. No entanto, se cada tarefa demora ou processa uma grande quantidade de dados, você pode optar por não reutilizar o JVM para garantir que toda a memória esteja liberada para as tarefas subsequentes. Ao usar uma versão da AMI, você pode personalizar o fator de replicação usando a ação de bootstrap `configure-hadoop` para definir a propriedade `mapred.job.reuse.jvm.num.tasks`.

O exemplo a seguir demonstra como configurar o fator de reutilização da JVM para a reutilização infinita da JVM.

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr create-cluster --name "Test cluster" --ami-version 3.11.0 \
--applications Name=Hue Name=Hive Name=Pig \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-groups InstanceGroupType=MASTER,InstanceCount=1,InstanceType=m3.xlarge \
InstanceGroupType=CORE,InstanceCount=2,InstanceType=m3.xlarge \
--bootstrap-actions Path=s3://elasticmapreduce/bootstrap-actions/configure-hadoop,\
Name="Configuring infinite JVM reuse",Args=["-m","mapred.job.reuse.jvm.num.tasks=-1"]
```

Detalhes da aplicação Hive para versões anteriores da AMI do Amazon EMR

Arquivos de log

Com o uso das versões 2.x e 3.x da AMI do Amazon EMR, os logs do Hive são salvos em `/mnt/var/log/apps/`. Para oferecer suporte a versões simultâneas do Hive, a versão do Hive que você executa determina o nome do arquivo de log, conforme mostrado na tabela a seguir.

Versão do Hive	Nome do arquivo de log
0.13.1	`hive.log` **Note** Desde esta desta versão, o Amazon EMR usa um nome de arquivo sem versionamento, `hive.log`. Versões secundárias compartilham o mesmo local de log que a versão principal.
0.11.0	`hive_0110.log` **Note** As versões secundárias do Hive 0.11.0, como o Hive 0.11.0.1, compartilham o mesmo local de arquivo de log que o Hive 0.11.0.
0.8.1	`hive_081.log` **Note** As versões secundárias do Hive 0.8.1, como o Hive 0.8.1.1, compartilham o mesmo local de arquivo de log que o Hive 0.8.1.
0.7.1	`hive_07_1.log`

Versão do Hive	Nome do arquivo de log
	Note As versões secundárias do Hive 0.7.1, como o Hive 0.7.1.3 e o Hive 0.7.1.4, compartilham o mesmo local de arquivo de log que o Hive 0.7.1.
0.7	`hive_07.log`
0,5	`hive_05.log`
0.4	`hive.log`

Funcionalidade de entrada dividida

Para implementar a funcionalidade de entrada dividida usando as versões do Hive anteriores à versão 0.13.1 (versões da AMI do Amazon EMR anteriores à versão 3.11.0), use o seguinte:

```
hive> set hive.input.format=org.apache.hadoop.hive.ql.io.HiveCombineSplitsInputFormat;
hive> set mapred.min.split.size=100000000;
```

Essa funcionalidade se tornou obsoleta com o Hive 0.13.1. Para obter a mesma funcionalidade de formato de entrada dividido na versão 3.11.0 da AMI do Amazon EMR, use o seguinte:

```
set hive.hadoop.supports.splittable.combineinputformat=true;
```

Portas de serviço do Thrift

O Thrift é uma estrutura RPC que define um formato compacto de serialização de binários usado para manter estruturas de dados para análise posterior. Normalmente, o Hive configura o servidor para operar nas seguintes portas.

Versão do Hive	Número da porta
Hive 0.13.1	10000
Hive 0.11.0	10004
Hive 0.8.1	10003
Hive 0.7.1	10002
Hive 0.7	10001
Hive 0.5	10000

Para obter mais informações sobre serviços do Thrift, consulte http://wiki.apache.org/thrift/.

Usar o Hive para recuperar partições

O Amazon EMR inclui uma instrução na linguagem de consulta Hive que recupera as partições de uma tabela dos dados de tabela localizados no Amazon s3. O exemplo a seguir mostra isso.

```
CREATE EXTERNAL TABLE (json string) raw_impression
PARTITIONED BY (dt string)
LOCATION 's3://elastic-mapreduce/samples/hive-ads/tables/impressions';
ALTER TABLE logs RECOVER PARTITIONS;
```

Os dados e diretórios da partição devem estar na localização especificada na definição de tabela e devem ser nomeados de acordo com a convenção do Hive: por exemplo, `dt=2009-01-01`.

> **Note**
>
> Depois do Hive 0.13.1, esse recurso tem suporte nativamente usando `msck repair table` e, portanto, `recover partitions` não tem suporte. Para obter mais informações, consulte https://cwiki.apache. org/confluence/display/Hive/LanguageManual+DDL.

Transmitir uma variável do Hive para um script

Para transmitir uma variável em uma etapa do Hive usando a AWS CLI, digite o comando a seguir, *myKey* substitua pelo nome do par de EC2 chaves e *amzn-s3-demo-bucket* substitua pelo nome do bucket. Neste exemplo, `SAMPLE` é um valor de variável precedido pela opção `-d`. Essa variável é definida no script Hive como: `${SAMPLE}`.

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr create-cluster --name "Test cluster" --ami-version 3.9 \
--applications Name=Hue Name=Hive Name=Pig \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-type m3.xlarge --instance-count 3 \
--steps Type=Hive,Name="Hive Program",ActionOnFailure=CONTINUE,\
Args=[-f,s3://elasticmapreduce/samples/hive-ads/libs/response-time-stats.q,-d,\
INPUT=s3://elasticmapreduce/samples/hive-ads/tables,-d,OUTPUT=s3://amzn-s3-demo-bucket/
hive-ads/output/,\
-d,SAMPLE=s3://elasticmapreduce/samples/hive-ads/]
```

Especificar um local de metastore externo

O procedimento a seguir mostra como substituir os valores de configuração padrão para a localização da metastore do Hive e iniciar um cluster usando a localização da metastore reconfigurada.

Para criar uma metastore localizada fora do cluster do EMR

1. Crie um banco de dados MySQL ou Aurora usando o Amazon RDS.

   Para obter informações sobre como criar um banco de dados do Amazon RDS, consulte Conceitos básicos do Amazon RDS.
2. Modifique seus grupos de segurança para permitir conexões JDBC entre seu banco de dados e o grupo de segurança ElasticMapReduce-Master.

Para obter informações sobre como modificar os grupos de segurança quanto ao acesso, consulte Grupos de segurança do Amazon RDS no Guia do usuário do Amazon RDS.

3. Defina os valores de configuração do JDBC em `hive-site.xml`:

   a. Crie um arquivo de configuração `hive-site.xml` contendo as seguintes informações:

```
<configuration>
  <property>
    <name>javax.jdo.option.ConnectionURL</name>
    <value>jdbc:mariadb://hostname:3306/hive?createDatabaseIfNotExist=true</
value>
    <description>JDBC connect string for a JDBC metastore</description>
  </property>
  <property>
    <name>javax.jdo.option.ConnectionUserName</name>
    <value>hive</value>
    <description>Username to use against metastore database</description>
  </property>
  <property>
    <name>javax.jdo.option.ConnectionPassword</name>
    <value>password</value>
    <description>Password to use against metastore database</description>
  </property>
</configuration>
```

   *hostname*é o endereço DNS da instância do Amazon RDS que executa o banco de dados. *username*e *password* são as credenciais do seu banco de dados. Para obter mais informações sobre como se conectar a instâncias de banco de dados MySQL e Aurora, consulte Conectar-se a uma instância de banco de dados executando o mecanismo de banco de dados MySQL e Conectar-se a um cluster de banco de dados do Aurora no Guia do usuário do Amazon RDS.

   Os drivers JDBC são instalados pelo Amazon EMR.

   > **Note**
   >
   > A propriedade "value" não deve conter espaços ou retornos de carro. Eles devem aparecer todos em uma única linha.

b. Salve o arquivo `hive-site.xml` em um local no Amazon S3, como `s3://amzn-s3-demo-bucket/hive-site.xml`.

4. Crie um cluster especificando a localização do Amazon S3 do arquivo `hive-site.xml` personalizado.

   O exemplo de comando a seguir demonstra um AWS CLI comando que faz isso.

   > **Note**
   >
   > Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr create-cluster --name "Test cluster" --ami-version 3.10 \
--applications Name=Hue Name=Hive Name=Pig \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-type m3.xlarge --instance-count 3 \
--bootstrap-actions Name="Install Hive Site Configuration",\
Path="s3://region.elasticmapreduce/libs/hive/hive-script",\
Args=["--base-path","s3://elasticmapreduce/libs/hive","--install-hive-site",\
"--hive-site=s3://amzn-s3-demo-bucket/hive-site.xml","--hive-versions","latest"]
```

Conectar-se com o Hive usando JDBC

Para se conectar ao Hive por meio do JDBC você precisa fazer o download do driver do JDBC e instalar um cliente SQL. O exemplo a seguir demonstra o uso do SQL Workbench/J para estabelecer a conexão com o Hive usando o JDBC.

Para fazer download dos drivers do JDBC

1. Faça o download e extraia os drivers apropriados para as versões do Hive que você deseja acessar. A versão do Hive depende da AMI escolhida na criação de um cluster do Amazon EMR.

   - Drivers JDBC Hive 0.13.1: JDBC_1.0.4.1004.zip https://amazon-odbc-jdbc-drivers.s3.amazonaws.com/public/ AmazonHive

- Drivers JDBC do Hive 0.11.0: https://mvnrepository.com/artifact/org.apache.hive/hive-jdbc/0.11.0
- Drivers JDBC do Hive 0.8.1: https://mvnrepository.com/artifact/org.apache.hive/hive-jdbc/0.8.1

2. Instale o SQL Workbench/J. Para obter mais informações, consulte Instalar e iniciar o SQL Workbench/J (apenas em inglês) no SQL Workbench/J User's Manual.
3. Crie um túnel do SSH no nó principal do cluster. A porta de conexão é diferente dependendo da versão do Hive. Comandos de exemplo são fornecidos nas tabelas abaixo para usuários do `ssh` do Linux e comandos do PuTTY para usuários do Windows

   Comandos SSH do Linux

Versão do Hive	Command
0.13.1	`ssh -o ServerAliveInterval=10 -i` *`path-to-key-file`* `-N -L 10000:localhost:10000 hadoop@` *`master-public-dns-name`*
0.11.0	`ssh -o ServerAliveInterval=10 -i` *`path-to-key-file`* `-N -L 10004:localhost:10004 hadoop@` *`master-public-dns-name`*
0.8.1	`ssh -o ServerAliveInterval=10 -i` *`path-to-key-file`* `-N -L 10003:localhost:10003 hadoop@` *`master-public-dns-name`*
0.7.1	`ssh -o ServerAliveInterval=10 -i` *`path-to-key-file`* `-N -L 10002:localhost:10002 hadoop@` *`master-public-dns-name`*
0.7	`ssh -o ServerAliveInterval=10 -i` *`path-to-key-file`* `-N -L 10001:localhost:10001 hadoop@` *`master-public-dns-name`*
0,5	`ssh -o ServerAliveInterval=10 -i` *`path-to-key-file`* `-N -L 10000:localhost:10000 hadoop@` *`master-public-dns-name`*

   Configurações do túnel no PuTTY do Windows

Versão do Hive	Configurações do túnel
0.13.1	Porta de origem: 10000 Destino ::10000 *`master-public-dns-name`*

Versão do Hive	Configurações do túnel
0.11.0	Porta de origem: 10004 Destino ::10004 *master-public-dns-name*
0.8.1	Porta de origem: 10003 Destino ::10003 *master-public-dns-name*

4. Adicione o driver do JDBC ao SQL Workbench.

   a. Na caixa de diálogo Select Connection Profile (Selecionar perfil de conexão), escolha Manage Drivers (Gerenciar drivers).

   b. Escolha o ícone Create a new entry (Criar uma nova entrada) (página em branco).

   c. No campo Name (Nome), digite **`Hive JDBC`**.

   d. Em Library (Biblioteca), clique no ícone Select the JAR file(s) (Selecionar os arquivos JAR).

   e. Selecione arquivos JAR, conforme mostrado na tabela a seguir.

Versão do driver do Hive	Arquivos JAR para adicionar

0.13.1	`hive_metastore.jar`
	`hive_service.jar`
	`HiveJDBC3.jar`
	`libfb303-0.9.0.jar`
	`libthrift-0.9.0.jar`
	`log4j-1.2.14.jar`
	`ql.jar`
	`slf4j-api-1.5.8.jar`
	`slf4j-log4j12-1.5.8.jar`
	`TCLIServiceClient.jar`

0.11.0	`hadoop-core-1.0.3.jar`
	`hive-exec-0.11.0.jar`
	`hive-jdbc-0.11.0.jar`
	`hive-metastore-0.11.0.jar`
	`hive-service-0.11.0.jar`
	`libfb303-0.9.0.jar`
	`commons-logging-1.0.4.jar`

Versão do driver do Hive	Arquivos JAR para adicionar
	slf4j-api-1.6.1.jar
0.8.1	hadoop-core-0.20.205.jar hive-exec-0.8.1.jar hive-jdbc-0.8.1.jar hive-metastore-0.8.1.jar hive-service-0.8.1.jar libfb303-0.7.0.jar libthrift-0.7.0.jar log4j-1.2.15.jar slf4j-api-1.6.1.jar slf4j-log4j12-1.6.1.jar
0.7.1	hadoop-0.20-core.jar hive-exec-0.7.1.jar hive-jdbc-0.7.1.jar hive-metastore-0.7.1.jar hive-service-0.7.1.jar libfb303.jar commons-logging-1.0.4.jar slf4j-api-1.6.1.jar slf4j-log4j12-1.6.1.jar
0.7	hadoop-0.20-core.jar hive-exec-0.7.0.jar hive-jdbc-0.7.0.jar hive-metastore-0.7.0.jar hive-service-0.7.0.jar libfb303.jar commons-logging-1.0.4.jar slf4j-api-1.5.6.jar slf4j-log4j12-1.5.6.jar

Versão do driver do Hive	Arquivos JAR para adicionar

0,5	`hadoop-0.20-core.jar`
	`hive-exec-0.5.0.jar`
	`hive-jdbc-0.5.0.jar`
	`hive-metastore-0.5.0.jar`
	`hive-service-0.5.0.jar`
	`libfb303.jar`
	`log4j-1.2.15.jar`
	`commons-logging-1.0.4.jar`

f. Na caixa de diálogo Please select one driver (Selecione um driver), selecione um driver de acordo com a tabela a seguir e clique em OK.

Versão do Hive	Nome da classe do driver
0.13.1	`com.amazon.hive.jdbc3.HS2Driver`
0.11.0	`org.apache.hadoop.hive.jdbc.HiveDriver.jar`
0.8.1	`org.apache.hadoop.hive.jdbc.HiveDriver.jar`
0.7.1	`org.apache.hadoop.hive.jdbc.HiveDriver.jar`
0.7	`org.apache.hadoop.hive.jdbc.HiveDriver.jar`
0,5	`org.apache.hadoop.hive.jdbc.HiveDriver.jar`

5. Quando você retornar para a caixa de diálogo Select Connection Profile (Selecionar perfil de conexão), verifique se o campo Driver está definido como Hive JDBC (JDBC do Hive) e forneça a string de conexão JDBC no campo URL de acordo com a tabela a seguir.

Versão do Hive	String de conexão do JDBC
0.13.1	`jdbc:hive2://localhost:10000/default`
0.11.0	`jdbc:hive://localhost:10004/default`
0.8.1	`jdbc:hive://localhost:10003/default`

Se o seu cluster usa a AMI versão 3.3.1 ou posterior, na caixa de diálogo Select Connection Profile (Selecionar perfil de conexão), digite **hadoop** no campo Username (Nome de usuário).

HBase detalhes da aplicação para versões anteriores da AMI do Amazon EMR

HBase Versões suportadas

HBase versão	AMI version	AWS CLI parâmetros de configuração	HBase detalhes da versão

0.94.18	3.1.0 e posterior	`--ami-version 3.1`
		`--ami-version 3.2`
		`--ami-version 3.3`
		`--applications Name=HBase`
			• Correções de bugs e melhorias.

0.94.7	3.0-3.0.4	`--ami-version 3.0`	

HBase versão	AMI version	AWS CLI parâmetros de configuração	HBase detalhes da versão
		`--applications Name=HBase`	

0.92	2.2 e posterior	`--ami-version 2.2 or later`
		
		`--applications Name=HBase`	

HBase pré-requisitos de clusters

Um cluster criado usando as versões 2.x e 3.x da AMI do Amazon EMR deve atender aos requisitos a seguir para. HBase

- O AWS CLI (opcional) — Para interagir HBase usando a linha de comando, baixe e instale a versão mais recente do AWS CLI. Para obter mais informações, consulte Instalar a AWS Command Line Interface no Guia do usuário da AWS Command Line Interface .
- Pelo menos duas instâncias (opcionais): o nó principal do cluster executa o servidor HBase mestre e o Zookeeper, enquanto os nós de tarefa executam os servidores da HBase região. Para obter o melhor desempenho, HBase os clusters devem ser executados em pelo menos duas EC2 instâncias, mas você pode ser HBase executado em um único nó para fins de avaliação.
- Cluster de longa execução: é executado HBase apenas em clusters de execução prolongada. Por padrão, a CLI e o console do Amazon EMR criam clusters de execução prolongada.
- Um par de EC2 chaves da Amazon key pair definido (recomendado): para usar o protocolo de rede Secure Shell (SSH) para conectar-se ao nó principal e executar comandos de HBase shell, você deve usar um par de EC2 chaves da Amazon ao criar o cluster.
- As versões corretas da AMI e do Hadoop — atualmente, os HBase clusters são suportados somente no Hadoop 20.205 ou posterior.
- Ganglia (opcional) — Para monitorar as métricas de HBase desempenho, instale o Ganglia ao criar o cluster.
- Um bucket do Amazon S3 para logs (opcional): os logs para HBase estão disponíveis no nó principal. Se você quiser que esses logs sejam copiados para o Amazon S3, especifique um bucket do S3 para receber arquivos de log quando criar o cluster.

Criação de um cluster com HBase

A tabela a seguir lista as opções que estão disponíveis ao ser usado o console para criar um cluster HBase usando uma versão da AMI do Amazon EMR.

Campo	Ação
Restore from backup (Restauração de backup)	Especifique se deseja pré-carregar o HBase cluster com dados armazenados no Amazon S3.
Backup location (Localização do backup)	Especifique o URI no qual o backup para restauração reside no Amazon S3.
Backup version (Versão do backup)	Opcionalmente, especifique o nome da versão do backup para uso em Backup Location (Localização do backup). Se você deixar esse campo em branco, o Amazon EMR usará o backup mais recente em Local de backup para preencher o novo cluster. HBase
Schedule Regular Backups (Agendar backups regulares)	Especifique se deseja agendar backups incrementais automáticos. O primeiro backup é um backup completo para criar uma linha de base para futuros backups incrementais.
Consistent backup (Backup consistente)	Especifica se os backups devem ser consistentes. Um backup consistente é aquele que pausa operações de gravação durante o estágio de backup inicial, a sincronização entre nós. Qualquer operações de gravação assim pausada é colocada em uma fila e retomada quando a sincronização é concluída.
Backup frequency (Frequência de backup)	O número de days/hours/minutes backups programados.

Campo	Ação
Backup location (Localização do backup)	O URI do Amazon S3 em que os backups são armazenados. O local de backup de cada HBase cluster deve ser diferente para garantir que os backups diferenciais permaneçam corretos.
Backup start time (Hora de início do backup)	Especifique quando o primeiro backup deve ocorrer. Você pode definir isso como now, o que faz com que o primeiro backup comece assim que o cluster estiver em execução, ou inserir uma data e uma hora no formato ISO. Por exemplo, 2012-06-15T20:00Z define a hora de início como 15 de junho de 2012 às 8 p.m. UTC.

O AWS CLI comando de exemplo a seguir inicia um cluster com HBase e outros aplicativos:

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr create-cluster --name "Test cluster" --ami-version 3.3 \
               --applications Name=Hue Name=Hive Name=Pig Name=HBase \
               --use-default-roles --ec2-attributes KeyName=myKey \
               --instance-type c1.xlarge --instance-count 3 --termination-protected
```

Depois que a conexão entre o Hive e os HBase clusters for estabelecida (conforme mostrado no procedimento anterior), você poderá acessar os dados armazenados no HBase cluster criando uma tabela externa no Hive.

O exemplo a seguir, quando executado no prompt do Hive, cria uma tabela externa que faz referência a dados armazenados em uma HBase tabela chamada`inputTable`. Em seguida, você pode fazer referência `inputTable` nas instruções do Hive para consultar e modificar os dados armazenados no HBase cluster.

> **Note**
>
> O exemplo a seguir usa protobuf-java-2.4.0a.jar na AMI 2.3.3, mas você deve modificar o exemplo para corresponder à sua versão. Para verificar qual é a sua versão do Protocol Buffers JAR, execute o comando no prompt de comando do Hive: `! ls /home/hadoop/lib;`.

```
add jar lib/emr-metrics-1.0.jar ;
               add jar lib/protobuf-java-2.4.0a.jar ;

               set hbase.zookeeper.quorum=ec2-107-21-163-157.compute-1.amazonaws.com ;

               create external table inputTable (key string, value string)
                    stored by 'org.apache.hadoop.hive.hbase.HBaseStorageHandler'
                     with serdeproperties ("hbase.columns.mapping" = ":key,f1:col1")
                     tblproperties ("hbase.table.name" = "t1");

               select count(*) from inputTable ;
```

Personalizando a configuração HBase

Embora as configurações padrão devam funcionar com a maioria das aplicações, você tem a flexibilidade de modificar suas definições HBase de configuração. Para fazer isso, execute um dos dois scripts de ação de bootstrap:

- configure-hbase-daemons— Configura propriedades dos daemons principal, regionserver e zookeeper. Essas propriedades incluem o tamanho da pilha e as opções a serem passadas para a Java Virtual Machine (JVM) quando o HBase daemon é iniciado. Você pode definir essas propriedades como argumentos na ação de bootstrap. Essa ação de bootstrap modifica o arquivo de configuração/home/hadoop/conf/hbase-user-env.sh no HBase cluster.
- configure-hbase: define configurações HBase específicas de local, como a porta à qual o HBase mestre deve ser vinculado e o número máximo de vezes que o cliente de CLI deve tentar novamente uma ação. Você pode definir essas configurações one-by-one, como argumentos na ação de bootstrap, ou pode especificar o local de um arquivo de configuração XML no Amazon S3. Essa ação de bootstrap modifica o arquivo de configuração the /home/hadoop/conf/hbase -site.xml no HBase cluster.

> **Note**
>
> Esses scripts, como outras ações de bootstrap, só podem ser executados quando o cluster é criado; você não pode usá-los para alterar a configuração de um HBase cluster em execução no momento.

Quando você executa as ações configure-hbase ou configure-hbase-daemonsbootstrap, os valores especificados substituem os valores padrão. Todos os valores não explicitamente definidos recebem os valores padrão.

Configurar HBase com essas ações de bootstrap é semelhante ao uso de ações de bootstrap no Amazon EMR para definir configurações do Hadoop e propriedades de daemons do Hadoop. A diferença é que HBase não tem opções de memória por processo. Em vez disso, as opções de memória são definidas usando o `--`*`daemon`*`-opts` argumento, onde *`daemon`* é substituído pelo nome do daemon a ser configurado.

Configurar HBase daemons

O Amazon EMR fornece uma ação de bootstrap,`s3://`*`region`*`.elasticmapreduce/bootstrap-actions/configure-hbase-daemons`, que você pode usar para alterar a configuração de HBase daemons, em que *`region`* é a região em que o cluster do cluster. HBase

Para configurar HBase daemons usando o AWS CLI, adicione a ação `configure-hbase-daemons` bootstrap ao iniciar o cluster para configurar um ou mais daemons. HBase É possível definir as seguintes propriedades.

Propriedade	Descrição
`hbase-master-opts`	Opções que controlam como a JVM executa o daemon principal. Se definidas, estas substituirão as variáveis HBASE_MASTER_OPTS padrão.
`regionserver-opts`	Opções que controlam como a JVM executa o daemon do servidor de regiões. Se definidas, estas substituirão as variáveis HBASE_REGIONSERVER_OPTS padrão.

Propriedade	Descrição
`zookeeper-opts`	Opções que controlam como a JVM executa o daemon zookeeper. Se definidas, estas substituirão as variáveis HBASE_ZOOKEEPER_OPTS padrão.

Para obter mais informações sobre essas opções, consulte hbase-env.sh na HBase documentação.

O exemplo a seguir mostra uma ação de bootstrap para configurar valores para `zookeeper-opts` e `hbase-master-opts`.

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr create-cluster --name "Test cluster" --ami-version 3.3 \
--applications Name=Hue Name=Hive Name=Pig Name=HBase \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-type c1.xlarge --instance-count 3 --termination-protected \
--bootstrap-actions Path=s3://elasticmapreduce/bootstrap-actions/configure-hbase-daemons,\
Args=["--hbase-zookeeper-opts=-Xmx1024m -XX:GCTimeRatio=19","--hbase-master-opts=-Xmx2048m","--hbase-regionserver-opts=-Xmx4096m"]
```

Definir configurações de HBase local

O Amazon EMR fornece uma ação de bootstrap,`s3://elasticmapreduce/bootstrap-actions/configure-hbase`, que você pode usar para alterar a configuração do. HBase Você pode definir valores de configuração one-by-one, como argumentos na ação de bootstrap, ou pode especificar o local de um arquivo de configuração XML no Amazon S3. Definir valores de configuração one-by-one é útil se você precisar definir apenas algumas configurações. Defini-los usando um arquivo XML é útil quando há muitas mudanças por fazer, ou se você deseja salvar as definições de configuração para reutilização.

> **Note**
>
> Você pode acrescentar ao nome do bucket do Amazon S3 um prefixo de região, `s3://region.elasticmapreduce/bootstrap-actions/configure-hbase` como, por exemplo, em que região *region* em que o cluster do Amazon S3. HBase

Essa ação de bootstrap modifica o arquivo `/home/hadoop/conf/hbase-site.xml` de configuração no HBase cluster. A ação bootstrap só pode ser executada quando o HBase cluster é iniciado.

Para obter mais informações sobre as configurações de HBase local que você pode definir, consulte Configuração padrão na HBase documentação.

Defina a ação de `configure-hbase` bootstrap ao iniciar o HBase cluster e especifique os valores a `hbase-site.xml` serem alterados.

Para especificar configurações individuais HBase do site usando o AWS CLI

- Para alterar a `hbase.hregion.max.filesize` configuração, digite o comando a seguir e *myKey* substitua pelo nome do par de EC2 chaves da Amazon.

  > **Note**
  >
  > Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr create-cluster --name "Test cluster" --ami-version 3.3 \
--applications Name=Hue Name=Hive Name=Pig Name=HBase \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-type c1.xlarge --instance-count 3 --termination-protected \
--bootstrap-actions Path=s3://elasticmapreduce/bootstrap-actions/configure-
hbase,Args=["-s","hbase.hregion.max.filesize=52428800"]
```

Para especificar as configurações do HBase site com um arquivo XML usando o AWS CLI

1. Crie uma versão personalizada de `hbase-site.xml`. Seu arquivo personalizado deve ser um XML válido. Para reduzir as probabilidades da ocorrência de erros, comece com a cópia padrão de`hbase-site.xml`, localizada no nó HBase principal do Amazon EMR em`/home/hadoop/conf/hbase-site.xml`, e edite uma cópia desse arquivo em vez de criar um arquivo do zero. Você pode dar um novo nome para o seu novo arquivo ou manter `hbase-site.xml`.
2. Carregue seu arquivo `hbase-site.xml` personalizado em um bucket do Amazon S3. Ele deve ter permissões definidas para que a AWS conta da que inicia o cluster possa acessar o arquivo. Se a AWS conta da que iniciar o cluster também possuir o bucket do Amazon S3, ela terá acesso.
3. Defina a ação de bootstrap configure-hbase ao iniciar o HBase cluster e inclua a localização do seu arquivo personalizado. `hbase-site.xml` O exemplo a seguir define os valores de configuração do HBase site para aqueles especificados no arquivos`3://amzn-s3-demo-bucket/my-hbase-site.xml`. Digite o comando a seguir, *`myKey`* substitua pelo nome do par de EC2 chaves e *`amzn-s3-demo-bucket`* substitua pelo nome do bucket do Amazon S3.

   > **Note**
   >
   > Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

   ```
   aws emr create-cluster --name "Test cluster" --ami-version 3.3 \
           --applications Name=Hue Name=Hive Name=Pig Name=HBase \
           --use-default-roles --ec2-attributes KeyName=myKey \
           --instance-type c1.xlarge --instance-count 3 --termination-protected \
           --bootstrap-actions Path=s3://elasticmapreduce/bootstrap-actions/configure-
   hbase,Args=["--site-config-file","s3://amzn-s3-demo-bucket/config.xml"]
   ```

   Se você especificar mais de uma opção para personalizar a HBase operação, deverá prefixar cada par de valores-chave com uma `-s` opção, conforme mostrado no exemplo a seguir:

   ```
          --bootstrap-actions s3://elasticmapreduce/bootstrap-actions/configure-
   hbase,Args=["-s","zookeeper.session.timeout=60000"]
   ```

Com o proxy definido e a conexão SSH aberta, você pode exibir a HBase interface do usuário abrindo uma janela do navegador com http: //:60010/master-status*master-public-dns-name*, em que é *master-public-dns-name* o endereço DNS público do nó principal no cluster. HBase

Você pode exibir os HBase logs atuais usando o SSH para se conectar ao nó principal e navegando até o `mnt/var/log/hbase` diretório. Esses logs não estão disponíveis depois que o cluster é encerrado, a menos que você habilite o registro em log no Amazon S3 quando o cluster for iniciado.

Fazer o backup e a restauração HBase

O Amazon EMR fornece a capacidade de fazer backup HBase dos dados para o Amazon S3 manualmente ou em uma programação automatizada. Você pode executar backups completos e incrementais. Depois de ter uma versão de backup dos HBase dados, você pode restaurar essa versão em um HBase cluster. Você pode restaurar em um HBase cluster que esteja em execução no momento ou iniciar um novo cluster pré-preenchido com dados de backup.

Durante o processo de backup, HBase continua executando comandos de gravação. Embora isso garanta que o cluster permanecerá disponível durante o backup, existe o risco de inconsistência entre os dados que estão sendo copiados para backup e quaisquer operações de gravação que estejam sendo executadas em paralelo. Para entender as inconsistências que podem surgir, é preciso considerar que HBase distribui as operações de gravação entre os nós em seu cluster. Se uma operação de gravação acontecer depois que um nó específico for sondado, esses dados não serão incluídos no arquivamento de backup. Você pode até descobrir que gravações anteriores no HBase cluster (enviadas para um nó que já foi pesquisado) podem não estar no arquivo de backup, enquanto gravações posteriores (enviadas para um nó antes de ser pesquisado) estão incluídas.

Se for necessário um backup consistente, você deverá pausar as gravações HBase durante a parte inicial do processo de backup, a sincronização entre os nós. Isso pode ser feito especificando o parâmetro `--consistent` ao solicitar um backup. Com esse parâmetro, as gravações durante esse período são enfileiradas e executadas logo após a conclusão da sincronização. Você também pode agendar backups recorrentes, que resolvem quaisquer inconsistências com o passar do tempo, conforme os dados perdidos em uma passagem são copiados para backup na passagem seguinte.

Ao fazer backup HBase dos dados, você deve especificar um diretório de backup diferente para cada cluster. Uma maneira fácil de fazer isso é usar o identificador de cluster como parte do caminho especificado para o diretório de backup. Por exemplo, `.s3://amzn-s3-demo-bucket/backups/`

`j-3AEXXXXXX16F2` Isso garante que qualquer backup incremental futuro faça referência ao HBase cluster correto.

Quando você estiver pronto para excluir arquivos de backup antigos que não são mais necessários, recomendamos que você primeiro faça um backup completo dos seus HBase dados. Isso garante que todos os dados sejam preservados e fornece uma base para futuros backups incrementais. Terminado o backup completo, você pode navegar até o local do backup e excluir manualmente os arquivos de backup antigos.

O processo HBase de backup usa o S3 DistCp para a operação de cópia, que tem certas limitações em relação ao espaço de armazenamento temporário de arquivos.

Fazer o backup e a restauração HBase usando o console

O console oferece a capacidade de iniciar um novo cluster e preenchê-lo com dados de um HBase backup anterior. Ele também oferece a capacidade de programar backups incrementais periódicos dos HBase dados. Funcionalidades adicionais de backup e restauração, como a capacidade de restaurar dados para um cluster já em execução, fazer backups manuais e agendar backups completos automatizados, estão disponíveis usando a CLI.

Para preencher um novo cluster com HBase dados arquivados usando o console

1. Navegue até o novo console do Amazon EMR e selecione Alternar para o console antigo na navegação lateral. Para obter mais informações sobre o que esperar ao alternar para o console antigo, consulte Usar o console antigo.
2. Selecione Criar cluster.
3. Na seção Configuração de software, para Aplicativos adicionais, escolha Configurar HBasee adicionar.
4. Na caixa de diálogo Add Application (Adicionar aplicativo), marque Restore From Backup (Restaurar a partir do backup).
5. Em Local do Backup, especifique o local do backup a ser carregado no novo HBase cluster. Este deve ser um URL do Amazon S3; no formato `s3://amzn-s3-demo-bucket/backups/`.
6. Para Backup Version (Versão de backup), você tem a opção de especificar o nome de uma versão de backup a ser carregada, definindo um valor. Se você não definir um valor para Versão de backup, o Amazon EMR carregará o backup mais recente no local especificado.
7. Escolha Add (Adicionar) e continue para criar o cluster com outras opções, conforme desejado.

Para agendar backups automatizados de HBase dados usando o console

1. Na seção Configuração de software, para Aplicativos adicionais, escolha Configurar HBasee adicionar.
2. Escolha Schedule Regular Backups (Programar backups regulares).
3. Especifica se os backups devem ser consistentes. Um backup consistente é aquele que pausa operações de gravação durante o estágio de backup inicial, a sincronização entre nós. Qualquer operação de gravação pausada é colocada em uma fila e retomada quando a sincronização é concluída.
4. Defina com que frequência os backups devem ocorrer, inserindo um número para Backup Frequency (Frequência de backup) e escolhendo Days (Dias), Hours (Horas) ou Minutes (Minutos). O primeiro backup automatizado que é executado é um backup completo. Depois disso, o Amazon EMR salva backups incrementais, com base no horário que você especificar.
5. Especifique o local no Amazon S3 em que os backups devem ser armazenados. Cada HBase cluster deve ser copiado para backup em um local separado no Amazon S3 para garantir que os backups incrementais sejam calculados corretamente.
6. Especifique quando o primeiro backup deve ocorrer, definindo um valor para Backup Start Time (Hora de início do backup). Você pode definir isso como `now`, o que faz com que o primeiro backup comece assim que o cluster estiver em execução, ou inserir uma data e uma hora no formato ISO. Por exemplo, 2013-09-26T20:00Z, define a hora de início como 26 de setembro de 2013 às 8 p.m. UTC.
7. Escolha Adicionar.
8. Continue com a criação do cluster, com outras opções conforme desejado.

Monitor HBase com CloudWatch

O Amazon EMR informa três métricas CloudWatch que você pode usar para monitorar seus HBase backups. Essas métricas são enviadas CloudWatch em intervalos de cinco minutos e fornecidas sem custo.

Métrica	Descrição
`HBaseBackupFailed`	Se o último backup falhou. É definido como 0 por padrão e atualizado para 1 se a tentativa de backup anterior falhou. Essa métrica é relatada somente para HBase clusters.

Métrica	Descrição
	Caso de uso: monitorar HBase backups Unidades: Contagem
`HBaseMostRecentBackupDuration`	O tempo que o backup anterior levou para ser concluído. Essa métrica é definida independentemente de o último backup ter sido bem-sucedido ou ter falhado. Enquanto o backup está sendo processado, essa métrica retorna o número de minutos desde o início do backup. Essa métrica é relatada somente para HBase clusters. Caso de uso: monitorar HBase backups Unidade: minutos
`HBaseTimeSinceLastSuccessfulBackup`	O número de minutos decorridos após o início do último HBase backup bem-sucedido em seu cluster. Essa métrica é relatada somente para HBase clusters. Caso de uso: monitorar HBase backups Unidade: minutos

Configurar o Ganglia para HBase

Você configura o Ganglia para HBase usar a ação configure-hbase-for-gangliabootstrap. Essa ação de bootstrap é configurada HBase para publicar métricas no Ganglia.

Você deve configurar o HBase Ganglia ao iniciar o cluster; os relatórios do Ganglia não podem ser adicionados a um cluster em execução.

O Ganglia também armazena arquivos de log no servidor em `/mnt/var/log/ganglia/rrds`. Se você tiver configurado o cluster para manter a persistência dos arquivos de log em um bucket do Amazon S3, os arquivos de log do Ganglia serão mantidos lá também.

Para iniciar um cluster com o Ganglia for HBase, use a ação configure-hbase-for-gangliabootstrap conforme mostrado no exemplo a seguir.

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr create-cluster --name "Test cluster" --ami-version 3.3 \
--applications Name=Hue Name=Hive Name=Pig Name=HBase Name=Ganglia \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-type c1.xlarge --instance-count 3 --termination-protected \
--bootstrap-actions Path=s3://elasticmapreduce/bootstrap-actions/configure-hbase-for-ganglia
```

Depois que o cluster é executado com o Ganglia configurado, você pode acessar os gráficos e relatórios do Ganglia usando a interface gráfica em execução no nó principal.

Detalhes da aplicação Pig para versões anteriores da AMI do Amazon EMR

Versões compatíveis do Pig

A versão do Pig que você pode adicionar ao cluster depende da versão da AMI do Amazon EMR e da versão do Hadoop em uso. A tabela abaixo mostra quais versões do AMI e do Hadoop são compatíveis com as diferentes versões do Pig. Recomendamos que você use a versão mais recente do Pig para aproveitar as melhorias de desempenho e as novas funcionalidades.

Quando você usa a API para instalar o Pig, a versão padrão é usada, a menos que você especifique `--pig-versions` como argumento para a etapa que carrega o Pig no cluster durante a chamada para. RunJobFlow

Versão do Pig	AMI version	Parâmetros de configuração	Detalhes da versão do Pig

0.12.0
 	
Notas da versão
 	
Documentação
	3.1.0 e posterior	`--ami-version 3.1`
		
		`--ami-version 3.2`
			Adiciona suporte para:
			
			• Streaming UDFs sem implement ações de JVM

Versão do Pig	AMI version	Parâmetros de configuração	Detalhes da versão do Pig
		`--ami-version 3.3`	• Operadores ASSERT e IN • Expressão CASE • AvroStorage como uma função incorporada do Pig. • ParquetLoader e ParquetStorer como funções integradas • BigInteger e BigDecimal tipos
0.11.1.1 Notas da versão Documentação	2.2 e posterior	`--pig-versions 0.11.1.1` `--ami-version 2.2`	Melhora a performance do comando LOAD PigStorage se a entrada reside no Amazon S3.
0.11.1 Notas da versão Documentação	2.2 e posterior	`--pig-versions 0.11.1` `--ami-version 2.2`	Adiciona suporte para JDK 7, Hadoop 2, funções definidas pelo usuário do Groovy, SchemaTuple otimização, novos operadores e muito mais. Para obter mais informações, consulte Log de alterações do Pig 0.11.1.

Versão do Pig	AMI version	Parâmetros de configuração	Detalhes da versão do Pig
0.9.2.2 Notas da versão Documentação	2.2 e posterior	`--pig-versions 0.9.2.2` `--ami-version 2.2`	Adiciona suporte para Hadoop 1.0.3.
0.9.2.1 Notas da versão Documentação	2.2 e posterior	`--pig-versions 0.9.2.1` `--ami-version 2.2`	Adiciona suporte para MapR.
0.9.2 Notas da versão Documentação	2.2 e posterior	`--pig-versions 0.9.2` `--ami-version 2.2`	Inclui várias melhorias no desempenho e correções de erros. Para obter informaçõ es completas sobre as alterações do Pig 0.9.2, acesse o Log de alterações do Pig 0.9.2.
0.9.1 Notas da versão Documentação	2,0	`--pig-versions 0.9.1` `--ami-version 2.0`	
0.6 Notas da versão	1,0	`--pig-versions 0.6` `--ami-version 1.0`	

Versão do Pig	AMI version	Parâmetros de configuração	Detalhes da versão do Pig

0.3
 	
Notas da versão
	1,0	`--pig-versions 0.3`
		
		`--ami-version 1.0`	

Detalhes da versão do Pig

O Amazon EMR é compatível com determinadas versões do Pig que podem ter patches adicionais do Amazon EMR aplicados. Você pode configurar qual versão do Pig será executada em clusters do Amazon EMR. Para obter mais informações sobre como fazer isso, consulte Apache Pig. As seções a seguir descrevem diferentes versões do Pig e os patches aplicados às versões carregadas no Amazon EMR.

Patches do Pig

Esta seção descreve os patches personalizados aplicados nas versões do Pig disponíveis com o Amazon EMR.

Patches do Pig 0.11.1.1

A versão 0.11.1.1 do Pig do Pig do Amazon EMR é uma versão de manutenção que melhora a performance do comando LOAD PigStorage se a entrada reside no Amazon S3.

Patches do Pig 0.11.1

A versão 0.11.1 do pig do Amazon EMR contém todas as atualizações fornecidas pela Apache Software Foundation e os patches cumulativos do Amazon EMR do Pig versão 0.9.2.2. No entanto, não há novos patches específicos do Amazon EMR no Pig 0.11.1.

Patches do Pig 0.9.2

A Apache Pig 0.9.2 é uma versão de manutenção do Pig. A equipe do Amazon EMR aplicou os seguintes patches à versão 0.9.2 do Pig do Amazon EMR.

Patch	Descrição
PIG-1429	Adiciona o tipo de dados Boolean ao Pig como um tipo de dados de primeira classe. Para obter mais informações, acesse https://issues.apache. org/jira/browse/PIG-1429. Status: Confirmado Fixed in Apache Pig Version (Corrigido no Apache Pig versão): 0.10
PIG-1824	Suporte a módulos de importação na UDF do Jython. Para obter mais informações, acesse https://issues.apache. org/jira/browse/PIG-1824. Status: Confirmado Fixed in Apache Pig Version (Corrigido no Apache Pig versão): 0.10
PIG-2010	Pacote registrado JARs no cache distribuído. Para obter mais informações, acesse https://issues.apache. org/jira/browse/PIG-2010. Status: Confirmado Fixed in Apache Pig Version (Corrigido no Apache Pig versão): 0.11
PIG-2456	Adiciona um arquivo ~ /.pigbootup onde o usuário pode especificar instruções padrão do Pig. Para obter mais informações, acesse https://issues.apache. org/jira/browse/PIG-2456. Status: Confirmado Fixed in Apache Pig Version (Corrigido no Apache Pig versão): 0.11

Patch	Descrição
PIG-2623	Support ao uso de caminhos do Amazon S3 para se registrar. UDFs Para obter mais informações, acesse https://issues.apache. org/jira/browse/PIG-2623. Status: Confirmado Fixed in Apache Pig Version (Corrigido no Apache Pig versão): 0.10, 0.11

Patches do Pig 0.9.1

A equipe do Amazon EMR aplicou os seguintes patches à versão 0.9.1 do Pig do Amazon EMR.

Patch	Descrição
Oferece suporte a arquivos JAR e scripts do Pig em dfs	Adiciona suporte para a execução de scripts e o registro de arquivos JAR armazenados em HDFS, no Amazon S3 ou em outros sistemas de arquivos distribuídos. Para obter mais informações, acesse https://issues.apache. org/jira/browse/PIG-1505. Status: Confirmado Fixed in Apache Pig Version (Corrigido no Apache Pig versão) 0.8.0
Oferece suporte a vários sistemas de arquivos no Pig	Adiciona suporte a scripts do Pig para a leitura de dados de um sistema de arquivos e gravação em outro sistema. Para obter mais informações, acesse https://issues.apache. org/jira/browse/PIG-1564. Status: Não confirmado Fixed in Apache Pig Version (Corrigido no Apache Pig versão) n/a

Patch	Descrição
Adicionar data, hora e string do Piggybank UDFs	Adicione data e hora e string UDFs para oferecer suporte a scripts personalizados do Pig. Para obter mais informações, acesse https://issues.apache. org/jira/browse/PIG-1565. Status: Não confirmado Fixed in Apache Pig Version (Corrigido no Apache Pig versão) n/a

Clusters interativos e em lote no Pig

O Amazon EMR permite que você execute scripts do Pig em dois modos:

- Interativo
- Lote

Ao iniciar um cluster de longa execução usando o console ou o AWS CLI, você pode se conectar usando ssh o nó principal como usuário do Hadoop e usar o shell Grunt para desenvolver e executar seus scripts Pig de forma interativa. Usar o Pig no modo interativo permite que você revise o script do Pig com mais facilidade do que no modo em lote. Depois de revisar o script do Pig no modo interativo com êxito, você pode carregar o script para o Amazon S3 e executar o script no modo em lote na produção. Você também pode enviar comandos do Pig interativamente em um cluster em execução para analisar e transformar os dados conforme necessário.

No modo em lote, você carrega o script do Pig para o Amazon S3 e, em seguida, envia o trabalho para o cluster como uma etapa. As etapas do Pig podem ser enviadas para um cluster de execução prolongada ou um cluster temporário.

Detalhes de aplicações do Spark com versões anteriores da AMI do Amazon EMR

Usar o Spark interativamente ou no modo em lote

O Amazon EMR permite a execução de aplicações do Spark em dois modos:

- Interativo

- Lote

Ao iniciar um cluster de longa execução usando o console ou o AWS CLI, você pode se conectar usando SSH ao nó principal como usuário do Hadoop e usar o shell do Spark para desenvolver e executar seus aplicativos Spark de forma interativa. Usar o Spark interativamente permite criar protótipos ou testar aplicativos Spark com mais facilidade do que em um ambiente em lote. Depois de revisar com sucesso a aplicação do Spark no modo interativo, você pode colocar a aplicação do JAR ou o programa do Python do sistema de arquivos local no nó principal do cluster do Amazon S3. Em seguida, você pode enviar o aplicativo como um fluxo de trabalho em lote.

No modo em lote, carregue o script do Spark no Amazon S3 ou no sistema de arquivos do nó principal local e, em seguida, envie o trabalho ao cluster como uma etapa. As etapas do Spark podem ser enviadas a um cluster de longa execução ou a um cluster transitório.

Criar um cluster com o Spark instalado

Para executar um cluster com o Spark instalado usando o console

1. Navegue até o novo console do Amazon EMR e selecione Alternar para o console antigo na navegação lateral. Para obter mais informações sobre o que esperar ao alternar para o console antigo, consulte Usar o console antigo.
2. Selecione Criar cluster.
3. Para Software Configuration (Configuração do software), escolha a versão da AMI de que você precisa.
4. Para Applications to be installed (Aplicativos a serem instalados), escolha Spark na lista e depois escolha Configure and add (Configurar e adicionar).
5. Adicione argumentos para alterar a configuração do Spark conforme desejado. Para obter mais informações, consulte Configurar o Spark. Escolha Adicionar.
6. Selecione outras opções conforme necessário e escolha Create cluster (Criar cluster).

O seguinte exemplo mostra como criar um cluster com o Spark usando Java:

```
AmazonElasticMapReduceClient emr = new AmazonElasticMapReduceClient(credentials);
SupportedProductConfig sparkConfig = new SupportedProductConfig()
   .withName("Spark");

RunJobFlowRequest request = new RunJobFlowRequest()
```

```
    .withName("Spark Cluster")
    .withAmiVersion("3.11.0")
    .withNewSupportedProducts(sparkConfig)
    .withInstances(new JobFlowInstancesConfig()
     .withEc2KeyName("myKeyName")
     .withInstanceCount(1)
     .withKeepJobFlowAliveWhenNoSteps(true)
     .withMasterInstanceType("m3.xlarge")
     .withSlaveInstanceType("m3.xlarge")
    );
RunJobFlowResult result = emr.runJobFlow(request);
```

Configurar o Spark

Você configura o Spark ao criar um cluster executando a ação bootstrap localizada no awslabs/emr-bootstrap-actions/sparkrepositório no Github. Para saber quais argumentos a ação de bootstrap aceita, consulte o arquivo LEIAME nesse repositório. A ação de bootstrap configura propriedades no arquivo `$SPARK_CONF_DIR/spark-defaults.conf`. Para obter mais informações sobre configurações, consulte o tópico Configuração do Spark na documentação do Spark. Você pode substituir "latest" (mais recente) na URL a seguir pelo número da versão do Spark que você está instalando, por exemplo, `2.2.0` http://spark.apache.org/docs/latest/configuration.html.

Você também pode configurar o Spark dinamicamente no momento do envio de cada aplicativo. Uma configuração para maximizar automaticamente a alocação de recursos para um executor está disponível usando o arquivo de configuração `spark`. Para obter mais informações, consulte Substituir as definições de configuração padrão do Spark.

Alterar as configurações padrão do Spark

O seguinte exemplo mostra como criar um cluster com o `spark.executor.memory` definido como 2G usando a AWS CLI.

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr create-cluster --name "Spark cluster" --ami-version 3.11.0 \
```

```
--applications Name=Spark, Args=[-d,spark.executor.memory=2G] --ec2-attributes
 KeyName=myKey \
--instance-type m3.xlarge --instance-count 3 --use-default-roles
```

Enviar trabalhos para o Spark

Para enviar trabalhos a um cluster, use uma etapa para executar o script `spark-submit` no seu cluster do EMR. Adicione a etapa usando o `addJobFlowSteps` método em AmazonElasticMapReduceClient:

```
AWSCredentials credentials = new BasicAWSCredentials(accessKey, secretKey);
AmazonElasticMapReduceClient emr = new AmazonElasticMapReduceClient(credentials);
StepFactory stepFactory = new StepFactory();
AddJobFlowStepsRequest req = new AddJobFlowStepsRequest();
req.withJobFlowId("j-1K48XXXXXXHCB");

List<StepConfig> stepConfigs = new ArrayList<StepConfig>();

StepConfig sparkStep = new StepConfig()
 .withName("Spark Step")
 .withActionOnFailure("CONTINUE")
 .withHadoopJarStep(stepFactory.newScriptRunnerStep("/home/hadoop/spark/bin/spark-
submit","--class","org.apache.spark.examples.SparkPi","/home/hadoop/spark/lib/spark-
examples-1.3.1-hadoop2.4.0.jar","10"));

stepConfigs.add(sparkStep);
req.withSteps(stepConfigs);
AddJobFlowStepsResult result = emr.addJobFlowSteps(req);
```

Substituir as definições de configuração padrão do Spark

Você pode querer substituir valores de configuração padrão do Spark para cada aplicativo. Você pode fazer isso ao enviar aplicativos usando uma etapa, que transmite essencialmente opções para `spark-submit`. Por exemplo, você pode querer alterar a memória alocada para um processo de executor modificando `spark.executor.memory`. Você pode fornecer ao switch `--executor-memory` um argumento semelhante ao seguinte:

```
/home/hadoop/spark/bin/spark-submit --executor-memory 1g --class
 org.apache.spark.examples.SparkPi /home/hadoop/spark/lib/spark-examples*.jar 10
```

Da mesma forma, você pode ajustar `--executor-cores` e `--driver-memory`. Em uma etapa, você forneceria os seguintes argumentos para a etapa:

```
--executor-memory 1g --class org.apache.spark.examples.SparkPi /home/hadoop/spark/lib/
spark-examples*.jar 10
```

Você também pode ajustar as configurações que não possam não ter uma opção interna usando `--conf`. Para obter mais informações sobre outras configurações que são ajustáveis, consulte o tópico Dynamically loading Spark properties na documentação do Apache Spark.

Diferenças do DistCp utilitário S3 com versões anteriores da AMI do Amazon EMR

DistCp Versões do S3 compatíveis com o Amazon EMR

As seguintes versões do S3 são compatíveis com DistCp as versões da AMI do Amazon EMR. DistCpAs versões S3 posteriores à 1.0.7 são encontradas diretamente nos clusters. Use o JAR em `/home/hadoop/lib` para obter os recursos mais recentes.

Versão	Descrição	Data de lançamento
1.0.8	Adiciona as opções `--appendToLastFile` , `--requirePreviousManifest` e `--storageClass` .	3 de janeiro de 2014
1.0.7	Adiciona a opção `--s3ServerSideEncryption` .	2 de maio de 2013
1.0.6	Adiciona a opção `--s3Endpoint` .	6 de agosto de 2012
1.0.5	Melhora a capacidade de especificar qual versão do S3 DistCp deve ser executada.	27 de junho de 2012
1.0.4	Melhora a opção `--deleteOnSuccess` .	19 de junho de 2012
1.0.3	Adiciona suporte às opções `--numberFiles` e `--startingIndex` .	12 de junho de 2012

Versão	Descrição	Data de lançamento
1.0.2	Melhora a nomeação de arquivos ao usar grupos.	6 de junho de 2012
1.0.1	Lançamento inicial do S3DistCp.	19 de janeiro de 2012

Adicionar uma etapa de DistCp cópia do S3 a um cluster

Para adicionar uma etapa de DistCp cópia do S3 a um cluster em execução, digite o comando a seguir, *j-3GYXXXXXX9IOK* substitua pelo ID do cluster e *amzn-s3-demo-bucket* substitua pelo nome de bucket do Amazon S3.

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr add-steps --cluster-id j-3GYXXXXXX9IOK \
--steps Type=CUSTOM_JAR,Name="S3DistCp step",Jar=/home/hadoop/lib/emr-s3distcp-1.0.jar,
\
Args=["--s3Endpoint,s3-eu-west-1.amazonaws.com",\
"--src,s3://amzn-s3-demo-bucket/logs/j-3GYXXXXXX9IOJ/node/",\
"--dest,hdfs:///output",\
"--srcPattern,.*[a-zA-Z,]+"]
```

Example Carregar CloudFront os logs da Amazon no HDFS

Este exemplo carrega CloudFront logs da Amazon no HDFS ao adicionar uma etapa a um cluster em execução. No processo, ele muda o formato de compactação de Gzip (o padrão do CloudFront) para LZO. Isso é útil, pois dados compactados usando LZO podem ser divididos em vários mapas conforme são descompactados, portanto, você não precisa aguardar até que a compactação seja concluída, como ocorre com o Gzip. Isso proporciona uma melhor performance quando você analisa os dados usando o Amazon EMR. Esse exemplo também melhora a performance ao usar

a expressão regular especificada na opção `--groupBy` para combinar todos os logs de uma determinada hora em um único arquivo. Os clusters do Amazon EMR são mais eficientes quando processam alguns arquivos grandes com compactação LZO do que quando processam vários arquivos pequenos com compactação Gzip. Para dividir arquivos LZO, você deve indexá-los e usar a biblioteca de terceiros hadoop-lzo.

Para carregar CloudFront logs da Amazon no HDFS, digite o comando a seguir, *j-3GYXXXXXX9IOK* substitua pelo ID do cluster e *amzn-s3-demo-bucket* substitua pelo nome do bucket do Amazon S3.

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr add-steps --cluster-id j-3GYXXXXXX9IOK \
--steps Type=CUSTOM_JAR,Name="S3DistCp step",Jar=/home/hadoop/lib/emr-s3distcp-1.0.jar,
\
Args=["--src,s3://amzn-s3-demo-bucket/cf","--dest,hdfs:///local",\
"--groupBy,.*XABCD12345678.([0-9]+-[0-9]+-[0-9]+-[0-9]+).*",\
"--targetSize,128",
"--outputCodec,lzo","--deleteOnSuccess"]
```

Considere o caso no qual o exemplo anterior é executado com os seguintes arquivos de log do CloudFront .

```
s3://amzn-s3-demo-bucket/cf/XABCD12345678.2012-02-23-01.HLUS3JKx.gz
s3://amzn-s3-demo-bucket/cf/XABCD12345678.2012-02-23-01.I9CNAZrg.gz
s3://amzn-s3-demo-bucket/cf/XABCD12345678.2012-02-23-02.YRRwERSA.gz
s3://amzn-s3-demo-bucket/cf/XABCD12345678.2012-02-23-02.dshVLXFE.gz
s3://amzn-s3-demo-bucket/cf/XABCD12345678.2012-02-23-02.LpLfuShd.gz
```

O S3 DistCp copia, concatena e compacta os arquivos nos dois arquivos a seguir, onde o nome do arquivo é determinado pela correspondência feita pela expressão regular.

```
hdfs:///local/2012-02-23-01.lzo
hdfs:///local/2012-02-23-02.lzo
```

O que há de novo?

Esta página descreve as alterações e as funcionalidades disponíveis nas versões mais recentes do Amazon EMR 7.x, 6.x e 5.x.

Essas notas de versão também estão disponíveis nas páginas do Amazon EMR 7.9.0, Amazon EMR 6.15.0 e Amazon EMR 5.36.2, junto com as versões do aplicativo, as versões dos componentes e as classificações de configuração disponíveis para cada versão.

- Para obter notas de versões anteriores, consulte o Arquivo de notas de versão do Amazon EMR.
- Para receber atualizações quando uma nova versão do Amazon EMR estiver disponível, assine o RSS feed das notas de versão do Amazon EMR.

> **Note**
>
> Versões posteriores do Amazon EMR usam o AWS Signature Version 4 (SigV4) para autenticar solicitações no Amazon S3. Recomendamos o uso de uma versão do Amazon EMR que ofereça suporte ao SigV4 para que você possa acessar novos buckets do S3 e evitar a interrupção das suas workloads. Para obter mais informações e uma lista das versões do Amazon EMR compatíveis com o SigV4, consulte Amazon EMR e AWS Signature versão 4.

Amazon EMR 7.9.0 (versão mais recente da série 7.x)

Novas versões do Amazon EMR são disponibilizadas em diferentes regiões durante um período de vários anos, começando com a primeira região na data da versão inicial. A versão mais recente pode não estar disponível em sua região durante esse período.

As notas de lançamento a seguir incluem informações sobre a versão 7.9.0 do Amazon EMR.

Novos recursos

- Atualizações da aplicação: Amazon EMR 7.9.0 application upgrades include AWS SDK for Java 2.31.16/1.12.782, Delta 3.3.0-amzn-1, Hadoop 3.4.1, Hudi 0.15.0-amzn-6, Iceberg 1.7.1-amzn-2, Spark 3.5.5, and Zookeeper 3.9.3.

- Suporte CM-CMK para interface de usuário de aplicativo persistente — Esse recurso estende a criptografia SSE aplicada aos registros gerenciados no bucket de serviços do EMR para honrar a chave mestra personalizada gerenciada pelo cliente (CM-CMK). Esse recurso permite que os clientes gerenciem a acessibilidade dos registros após o encerramento dos clusters.

Problemas conhecidos

- Quando a opção de registro do EMR on EC2 é desativada, os registros não estão disponíveis após o encerramento do cluster e o aplicativo persistente não UIs conterá nenhum registro.

Alterações, melhorias e problemas resolvidos

- Depreciação do Oozie — Como o OSS Oozie entrou no modo sótão, a última versão a incluir o Apache Oozie é a versão 7.9 do Amazon EMR. O Apache Oozie será removido da versão 7.10 e versões superiores do Amazon EMR. Recomendamos que você use o MWAA como alternativa ao Oozie. Os clientes podem consultar o utilitário para GitHub migrar os fluxos oozie-to-airflow-emrde trabalho existentes do Oozie para o MWAA.
- Embalagem de contêineres do YARN — A partir da versão 7.9.0 do Amazon EMR, a política de empacotamento de contêineres agora está disponível para o programador de capacidade do YARN, que se baseia na política de posicionamento de vários nós do YARN.
- O hudi-cli.sh — O script hudi-cli.sh foi descontinuado no EMR versão 7.9.0 e superior. A versão 7.9.0 e superior do Amazon EMR usa .jar. hudi-cli-bundle
- Clusters EMR na AMI AL2 023 habilitada para FIPS — A partir do EMR 7.9, os clientes podem iniciar clusters EMR na AMI 023 habilitada para FIPS. AL2 Essa alteração atualiza as assinaturas digitais e os algoritmos de hash do EMR para garantir que todos os pacotes do EMR sigam as políticas criptográficas recomendadas para o FIPS 140-3. Nota: Isso não abrange o status de certificação FIPS do EMR AL2 ou 023 - consulte as perguntas frequentes do FIPS 023 AL2para obter as informações mais recentes sobre o status de validação do FIPS 140-3.
- Latência de etapas — Esta versão resolve um problema em que a latência de etapas aumentava ao usar criptografia em trânsito sem selecionar explicitamente o aplicativo Hadoop para instalação.
- Lançamento de cluster com volumes criptografados com EBS — Esta versão corrige um problema em que uma chamada KMS em volumes criptografados com EBS pode causar falha na inicialização de um cluster EMR.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250527.1	6.1.134-152.225.amzn2023	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein), China (Pequim), Ásia-Pacífico (Malásia), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Leste dos EUA (Norte da Virgínia), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Malásia), Europa (Londres), Ásia-Pacífico (Melbourne), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (AWS GovCloud AWS GovCloud Oregon), México (Central), Oeste dos EUA (Norte da Califórnia), Ásia-Pacífico

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv), Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.7.2 250331.(	6.1.131-143.221.am zn2023	18 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central), Ásia-Pacífico (Tóquio)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250303.(	6.1.129-138.220.am zn2023	27 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250218.1	6.1.128-136.201.amzn2023	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderabad), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2023.6.2 250211.(	6.1.127-135.201.am zn2023	26 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary), Ásia-Pací fico (Malásia), Ásia-Pací fico (Malásia), México (Central)

Amazon EMR 6.15.0 (versão mais recente da série 6.x)

Novas versões do Amazon EMR são disponibilizadas em diferentes regiões durante um período de vários anos, começando com a primeira região na data da versão inicial. A versão mais recente pode não estar disponível em sua região durante esse período.

As notas da versão a seguir incluem informações sobre a versão 6.15.0 do Amazon EMR. As alterações são referentes à versão 6.14.0. Para obter informações sobre o cronograma da versão, consulte o Log de alterações 6.15.0.

Novos recursos

- Atualizações da aplicação: Amazon EMR 6.15.0 application upgrades include Apache Hadoop 3.3.6, Apache Hudi 0.14.0-amzn-0, Iceberg 1.4.0-amzn-0, and Trino 426.
- Lançamentos mais rápidos para clusters do EMR que são executados em EC2 — Agora é até 35% mais rápido lançar um Amazon EMR em um cluster. EC2 Com essa melhoria, a maioria dos clientes pode lançar seus clusters em cinco minutos ou menos.
- CodeWhisperer para o EMR Studio — Agora você pode usar a Amazon com o CodeWhisperer Amazon EMR Studio para obter recomendações em tempo real à medida que você escreve código. JupyterLab CodeWhisperer pode concluir seus comentários, concluir linhas únicas de código, fazer line-by-line recomendações e gerar funções totalmente formadas.
- Tempos de reinicialização de trabalhos mais rápidos com o Flink: com o Amazon EMR 6.15.0 e superior, vários novos mecanismos estão disponíveis para o Apache Flink a fim de melhorar o

tempo de reinicialização de trabalhos durante as operações de recuperação ou escalabilidade de tarefas. Isso otimiza a velocidade de recuperação e reinicialização dos gráficos de execução para melhorar a estabilidade do trabalho.

- Controle de acesso detalhado e em nível de tabela para formatos de tabela aberta — Com o Amazon EMR 6.15.0 e superior, quando você executa trabalhos do Spark no Amazon EMR em EC2 clusters que acessam dados no AWS Glue Data Catalog, você pode AWS Lake Formation usar para aplicar permissões em nível de tabela, linha, coluna e célula em tabelas baseadas em Hudi, Iceberg ou Delta Lake.
- Atualização do Hadoop: o Amazon EMR 6.15.0 inclui uma atualização do Apache Hadoop para a versão 3.3.6. O Hadoop 3.3.6 era a versão mais recente na época da implantação do Amazon EMR 6.15, lançada pela Apache em junho de 2023. Versões anteriores do Amazon EMR (6.9.0 até 6.14.x) usavam o Hadoop 3.3.3.

  A atualização inclui centenas de melhorias e correções, além de recursos com parâmetros de nós de dados reconfiguráveis, a opção `DFSAdmin` para iniciar operações de reconfiguração em massa em todos os nós de dados ativos e uma API vetorizada que permite aos leitores exigentes especificar vários intervalos de leitura. O Hadoop 3.3.6 também adiciona suporte ao HDFS APIs e semântica ao seu registro de gravação antecipada (WAL), para que ele possa ser executado em outras implementações de sistemas de armazenamento. HBase Para obter mais informações, consulte os logs de alterações das versões 3.3.4, 3.3.5 e 3.3.6 na documentação do Apache Hadoop.
- Suporte ao AWS SDK for Java, versão 2 - Os aplicativos Amazon EMR 6.15.0 podem usar o SDK for AWS Java nas versões 1.12.569 ou 2.20.160 se o aplicativo suportar a versão 2. O AWS SDK for Java 2.x é uma grande reescrita da base de código da versão 1.x. Ele foi criado com base no Java 8+ e adiciona vários recursos frequentemente solicitados. Entre eles, suporte para E/S sem bloqueio e capacidade de conectar uma implementação HTTP diferente no runtime. Para obter mais informações, incluindo um Guia de migração do SDK para Java da v1 à v2, consulte o guia AWS SDK para Java, versão 2.

Alterações, melhorias e problemas resolvidos

- Para melhorar seus clusters de alta disponibilidade do EMR, essa versão permite a conectividade com daemons do Amazon EMR em um host local que usa endpoints IPv6.
- Esta versão habilita o TLS 1.2 para comunicação com ZooKeeper provisionados em todos os nós primários do seu cluster de alta disponibilidade.

- Esta versão melhora o gerenciamento dos arquivos de log de ZooKeeper transações que são mantidos nos nós primários para minimizar os cenários em que os arquivos de log ultrapassam os limites e interrompem as operações do cluster.
- Essa versão torna a comunicação entre nós mais resiliente para clusters de alta disponibilidade do EMR. Essa melhoria reduz a chance de falhas na ação de bootstrap ou na inicialização do cluster.
- O Tez no Amazon EMR 6.15.0 introduz configurações que você pode especificar para abrir de forma assíncrona as divisões de entrada em uma divisão agrupada do Tez. Isso resulta em uma performance mais rápida das consultas de leitura quando há um grande número de divisões de entrada em uma única divisão agrupada do Tez. Para obter mais informações, consulte a Abertura assíncrona da divisão do Tez.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Malásia), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			d), Leste dos EUA (Norte da Virgínia), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Malásia), Europa (Londres), Ásia-Pacífico (Melbourne), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (AWS GovCloud AWS GovCloud Oregon), México (Central) , Oeste dos EUA (Norte da Califórnia), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv), Ásia-Pacífico (Taipei), Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Hyderabad), Ásia-Pacífico (Jacarta) Arta), Ásia-Pacífico (Malásia), Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Malásia) , Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (EAU), Canadá (Central) , Oeste do Canadá

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Calgary), (Leste dos EUA), AWS GovCloud (Oeste dos EUA), China (AWS GovCloud Pequim), China (Ningxia), Israel (Tel Aviv), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Centro Leste (Emirados Árabes Unidos), Ásia-Pacífico (Melbourne),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Israel (Tel Aviv), Oeste do Canadá (Calgary), Europa (Zurique), Ásia-Pacífico (Malásia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 001.0	4.14.352	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 816.0	4.14.350	21 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 809.0	4.14.349	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 719.0	4.14.348	25 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 223.0	4.14.336	8 de março de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 131.0	4.14.336	14 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 124.0	4.14.336	7 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 109.0	4.14.334	24 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 218.0	4.14.330	2 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 206.0	4.14.330	22 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 116.0	4.14.328	11 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 101.0	4.14.327	13 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

Amazon EMR 5.36.2 (versão mais recente da série 5.x)

Novas versões do Amazon EMR são disponibilizadas em diferentes regiões durante um período de vários anos, começando com a primeira região na data da versão inicial. A versão mais recente pode não estar disponível em sua região durante esse período.

As notas da versão a seguir incluem informações sobre a versão 5.36.2 do Amazon EMR. As alterações são referentes à versão 5.36.1. Para obter informações sobre o cronograma da versão, consulte o log de alterações.

Alterações, melhorias e problemas resolvidos

- Essa versão melhora a lógica de redução vertical da escala do cluster para que o Amazon EMR não reduza a escala dos nós centrais abaixo da configuração do fator de replicação do HDFS para o cluster. Essa melhoria atende aos requisitos de redundância de dados e reduz a chance de uma operação de ajuste de escala ser interrompida.
- Esta versão adiciona um mecanismo de nova tentativa ao fluxo de trabalho de escalabilidade de clusters para clusters do EMR que executam o Presto ou o Trino. Essa melhoria reduz o risco de o redimensionamento do cluster ser executado indefinidamente devido a uma única operação de redimensionamento com falha. Ela também aprimora a utilização dos clusters, porque seu cluster aumenta e reduz a escala verticalmente com mais rapidez.
- Corrige um problema em que as operações de redução vertical da escala do cluster podem ficar paralisadas enquanto o Amazon EMR desativa normalmente um nó central e ele se torna não íntegro antes de ser totalmente desativado.
- Melhora a estabilidade de um nó em um cluster de alta disponibilidade com vários nós primários quando o Amazon EMR reinicia um único nó.

- Otimiza o gerenciamento de registros com o Amazon EMR em execução na Amazon. EC2 Como resultado, é possível ver uma pequena redução nos custos de armazenamento dos logs do cluster.
- Melhora o gerenciamento dos arquivos de log de ZooKeeper transações que são mantidos nos nós primários para minimizar os cenários em que os arquivos de log ultrapassam os limites e interrompem as operações do cluster.
- Corrige um bug raro que pode fazer com que um cluster de alta disponibilidade com vários nós primários falhe por não conseguir se comunicar com o ResourceManager Yarn.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Malásia), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací fico (Seul), Ásia-Pací

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			fico (Osaka), Ásia-Pací fico (Malásia), Europa (Londres), Ásia-Pacífico (Melbourne), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (AWS GovCloud AWS GovCloud Oregon), México (Central) , Oeste dos EUA (Norte da Califórnia), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv), Ásia-Pacífico (Taipei), Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Hyderabad), Ásia-Pacífico (Jacarta) Arta), Ásia-Pacífico (Malásia), Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Malásia) , Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (EAU), Canadá (Central) , Oeste do Canadá

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Calgary), (Leste dos EUA), AWS GovCloud (Oeste dos EUA), China (AWS GovCloud Pequim), China (Ningxia), Israel (Tel Aviv), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Centro Leste (Emirados Árabes Unidos), Ásia-Pacífico (Melbourne),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Israel (Tel Aviv), Oeste do Canadá (Calgary), Europa (Zurique)
2.0.2024 001.0	4.14.352	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 816.0	4.14.350	21 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 809.0	4.14.349	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 719.0	4.14.348	25 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 503.0	4.14.343	xxxxxx, 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

Amazon EMR e AWS Signature versão 4

As versões do Amazon EMR usam AWS Signature versão 4 (SigV4) para autenticar solicitações no Amazon S3. Os buckets criados no Amazon S3 depois de 24 de junho de 2020 não são compatíveis com solicitações assinadas pelo Signature Version 2 (SigV2). Os buckets criados em ou antes de 24 de junho de 2020 continuarão a oferecer suporte ao SigV2. Recomendamos a migração para uma versão do Amazon EMR que ofereça suporte ao SigV4 para que você possa acessar novos buckets do S3 e evitar a interrupção das suas workloads.

Se você usa aplicações incluídas no Amazon EMR, como Apache Spark, Apache Hive e Presto, não precisa alterar o código da aplicação para usar o SigV4. Se você usa aplicações personalizadas que não estão incluídas no Amazon EMR, talvez seja necessário atualizar seu código para usar o SigV4. Para obter mais informações, consulte Migração do Signature Version 2 para o Signature Version 4 no Guia do usuário do Amazon S3.

As seguintes versões do Amazon EMR são compatíveis com SigV4: emr-4.7.4, emr-4.8.5, emr-4.9.6, emr-4.10.1, emr-5.1.1, emr-5.2.3, emr-5.3.2, emr-5.4.1, emr-5.5.4, emr-5.6.1, emr-5.7.1, emr-5.8.3, emr-5.9.1, emr-5.10.1, emr-5.11.4, emr-5.12.3, emr-5.13.1, emr-5.14.2, emr-5.15.1, emr-5.16.1, emr-5.17.2, emr-5.18.1, emr-5.19.1, emr-5.20.1, emr-5.21.2, and emr-5.22.0 and higher. Todas as versões 6.x e 7.x são compatíveis com SigV4.

Abordagem para mitigar o CVE-2021-44228

> **Note**
>
> Para as versões 6.9.0 e posteriores do Amazon EMR, todos os componentes instalados pelo Amazon EMR que usam bibliotecas do Log4j usam o Log4j versão 2.17.1 ou posterior.

Amazon EMR em execução em EC2

O problema discutido no CVE-2021-44228 é relevante para as versões do núcleo do Apache Log4j entre 2.0.0 e 2.14.1 ao serem processadas entradas de fontes não confiáveis. Os clusters do Amazon EMR iniciados com as versões 5.x do Amazon EMR até 5.34.0 e as versões do EMR 6.x até o Amazon EMR 6.5.0 incluem estruturas de código aberto, como Apache Hive, Flink, HUDI, Presto e Trino, que usam essas versões do Apache Log4j. No entanto, muitos clientes usam as estruturas de código aberto instaladas nos seus clusters do Amazon EMR para processar e registrar em log entradas de fontes não confiáveis.

Recomendamos que você aplique a “Solução de ação de bootstrap do Amazon EMR para Log4j CVE-2021-44228” conforme descrito na seção a seguir. Essa solução também trata do CVE-2021-45046.

> **Note**
>
> Os scripts de ação de bootstrap para Amazon EMR foram atualizados em 7 de setembro de 2022 para incluir correções incrementais de bugs e melhorias para o Oozie. Se você usa o Oozie, deve aplicar a solução de ação de bootstrap atualizada do Amazon EMR descrita na seção a seguir.

Amazon EMR no EKS

Se você usa o Amazon EMR no EKS com a configuração padrão, o problema descrito no CVE-2021-44228 não afeta você e não é preciso aplicar a solução descrita na seção Solução de ação bootstrap do Amazon EMR para Log4j CVE-2021-44228 e CVE-2021-45046. Para o Amazon EMR no EKS, o runtime do Amazon EMR para Spark usa o Apache Log4j versão 1.2.17. Ao usar o Amazon EMR no EKS, você não deve alterar a definição padrão do componente `log4j.appender` como `log`.

Solução de ação bootstrap do Amazon EMR para Log4j CVE-2021-44228 e CVE-2021-45046

Essa solução fornece uma ação de bootstrap do Amazon EMR que deve ser aplicada nos clusters do Amazon EMR. Para cada versão do Amazon EMR, você encontrará um link para um script de ação de bootstrap abaixo. Para aplicar essa ação de bootstrap, você deve concluir as seguintes etapas:

1. Copie o script que corresponde à sua versão do Amazon EMR para um bucket do S3 local na sua Conta da AWS. Certifique-se de estar usando um script de bootstrap específico para sua versão do Amazon EMR.
2. Configure uma ação de bootstrap para que os clusters do EMR executem o script copiado para o bucket do S3 de acordo com as instruções descritas na documentação do EMR. Se você tiver outras ações de bootstrap configuradas para os clusters do EMR, certifique-se de que esse script esteja configurado como o primeiro script de ação de bootstrap a ser executado.
3. Encerre os clusters EMR existentes e inicie novos clusters com o script de ação bootstrap. AWS recomenda que você teste os scripts de bootstrap em seu ambiente de teste e valide seus aplicativos antes de aplicá-los ao seu ambiente de produção. Se você não estiver usando a

revisão mais recente de uma versão secundária do EMR (por exemplo, 6.3.0), deverá usar a revisão mais recente (por exemplo, 6.3.1) e, em seguida, aplicar a solução discutida acima.

CVE-2021-44228 e CVE-2021-45046: scripts de bootstrap para versões do Amazon EMR

Número da versão do Amazon EMR	Local do script	Data da versão do script
6.5.0	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-6.5.0-v2.sh`	24 de março de 2022
6.4.0	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-6.4.0-v2.sh`	24 de março de 2022
6.3.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-6.3.1-v2.sh`	24 de março de 2022
6.2.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-6.2.1-v2.sh`	24 de março de 2022
6.1.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-6.1.1-v2.sh`	14 de dezembro de 2021
6.0.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-6.0.1-v2.sh`	14 de dezembro de 2021

Número da versão do Amazon EMR	Local do script	Data da versão do script
5.34,0	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.34.0-v2.sh`	12 de dezembro de 2021
5.3.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.33.1-v2.sh`	12 de dezembro de 2021
5.32.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.32.1-v2.sh`	13 de dezembro de 2021
5.31.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.31.1-v2.sh`	13 de dezembro de 2021
5.30.2	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.30.2-v2.sh`	14 de dezembro de 2021
5.29.0	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.29.0-v2.sh`	14 de dezembro de 2021

Número da versão do Amazon EMR	Local do script	Data da versão do script
5.28.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.28.1-v2.sh`	15 de dezembro de 2021
5.27.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.27.1-v2.sh`	15 de dezembro de 2021
5.26.0	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.26.0-v2.sh`	15 de dezembro de 2021
5.25.0	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.25.0-v2.sh`	15 de dezembro de 2021
5.24.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.24.1-v2.sh`	15 de dezembro de 2021
5.23.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.23.1-v2.sh`	15 de dezembro de 2021

Número da versão do Amazon EMR	Local do script	Data da versão do script
5.22.0	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.22.0-v2.sh`	15 de dezembro de 2021
5.21.2	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.21.2-v2.sh`	15 de dezembro de 2021
5.20.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.20.1-v2.sh`	15 de dezembro de 2021
5.19.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.19.1-v2.sh`	15 de dezembro de 2021
5.18.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.18.1-v2.sh`	15 de dezembro de 2021
5.17.2	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.17.2-v2.sh`	15 de dezembro de 2021

Número da versão do Amazon EMR	Local do script	Data da versão do script
5.16.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.16.1-v2.sh`	15 de dezembro de 2021
5.15.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.15.1-v2.sh`	15 de dezembro de 2021
5.14.2	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.14.2-v2.sh`	15 de dezembro de 2021
5.13.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.13.1-v2.sh`	15 de dezembro de 2021
5.12.3	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.12.3-v2.sh`	15 de dezembro de 2021
5.11.4	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.11.4-v2.sh`	15 de dezembro de 2021

Número da versão do Amazon EMR	Local do script	Data da versão do script
5.10.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.10.1-v2.sh`	15 de dezembro de 2021
5.9.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.9.1-v2.sh`	15 de dezembro de 2021
5.8.3	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.8.3-v2.sh`	15 de dezembro de 2021
5.7.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.7.1-v2.sh`	15 de dezembro de 2021

Versão de lançamento do EMR	Revisão mais recente em dezembro de 2021
6.3.0	6.3.1
6.2.0	6.2.1
6.1.0	6.1.1
6.0.0	6.0.1
5.33.0	5.3.1
5.32.0	5.32.1

Versão de lançamento do EMR	Revisão mais recente em dezembro de 2021
5.31.0	5.31.1
5.30.0 ou 5.30.1	5.30.2
5.28.0	5.28.1
5.27.0	5.27.1
5.24,0	5.24.1
5.23,0	5.23.1
5.21.0 ou 5.21.1	5.21.2
5.20.0	5.20.1
5.19.0	5.19.1
5.18.0	5.18.1
5.17.0 ou 5.17.1	5.17.2
5.16.0	5.16.1
5.15.0	5.15.1
5.14.0 ou 5.14.1	5.14.2
5.13.0	5.13.1
5.12.0, 5.12.1, 5.12.2	5.12.3
5.11.0, 5.11.1, 5.11.2, 5.11.3	5.11.4
5.9.0	5.9.1
5.8.0, 5.8.1, 5.8.2	5.8.3
5.7.0	5.7.1

Perguntas frequentes

- As versões do EMR anteriores ao EMR 5 são afetadas pelo CVE-2021-44228?

  Não. As versões do EMR anteriores à versão 5 do EMR usam versões do Log4j anteriores à 2.0.

- Essa solução aborda o CVE-2021-45046?

  Sim, essa solução também aborda o CVE-2021-45046.

- A solução lida com aplicações personalizadas que instalo nos meus clusters do EMR?

  O script de bootstrap só atualiza os arquivos JAR que são instalados pelo EMR. Se você instalar e executar aplicações personalizadas e arquivos JAR nos clusters do EMR por meio de ações de bootstrap, como etapas enviadas aos clusters, usando a AMI personalizada do Amazon Linux ou por meio de qualquer outro mecanismo, trabalhe com o fornecedor da aplicação para determinar se suas aplicações personalizadas são afetadas pelo CVE-2021-44228 e determine uma solução apropriada.

- Como devo lidar com imagens do docker personalizadas com o EMR no EKS?

  Se você adicionar aplicativos personalizados ao Amazon EMR no EKS usando imagens docker personalizadas ou enviar trabalhos para o Amazon EMR em arquivos de aplicativos EKSwith personalizados, trabalhe com o fornecedor do aplicativo para determinar se seus aplicativos personalizados são afetados pelo CVE-2021-44228 e determine uma solução apropriada.

- Como o script de bootstrap funciona para mitigar o problema descrito no CVE-2021-44228 e no CVE-2021-45046?

  O script de bootstrap atualiza as instruções de inicialização do EMR ao adicionar um novo conjunto de instruções. Essas novas instruções excluem os arquivos de JndiLookup classe usados pelo Log4j por todas as estruturas de código aberto instaladas pelo EMR. Isso segue a recomendação publicada pela Apache para lidar com os problemas do Log4j.

- Existe alguma atualização do EMR que usa as versões 2.17.1 ou superiores do Log4j?

  As versões do EMR 5 até a versão 5.34 e as versões do EMR 6 até a versão 6.5 usam versões mais antigas de estruturas de código aberto que são incompatíveis com as versões mais recentes do Log4j. Se você continuar usando essas versões, recomendamos que você aplique a ação bootstrap para mitigar os problemas discutidos no. CVEs Após a versão 5.34 do EMR 5 e a versão 6.5 do EMR 6, as aplicações que usam o Log4j 1.x e o Log4j 2.x serão atualizadas para usar o

Log4j 1.2.17 (ou superior) e o Log4j 2.17.1 (ou superior), respectivamente, e não exigirão o uso das ações de bootstrap indicadas acima para mitigar os problemas do CVE.

- As versões do EMR são afetadas pelo CVE-2021-45105?

  As aplicações instaladas pelo Amazon EMR com as configurações padrão do EMR não são afetados pelo CVE-2021-45105. Entre as aplicações instaladas pelo Amazon EMR, somente o Apache Hive usa o Apache Log4j com pesquisas de contexto e não usa layout de modelo não padrão de uma forma que permita o processamento de dados de entrada inadequados.

- O Amazon EMR é afetado por alguma das divulgações do CVE a seguir?

  A tabela a seguir contém uma lista dos CVEs que estão relacionados ao Log4j e indica se cada CVE afeta o Amazon EMR. As informações dessa tabela só se aplicam quando as aplicações são instaladas pelo Amazon EMR usando as configurações padrão.

CVE	Afeta o EMR	Observações
CVE-2022-23302	Não	O Amazon EMR não configura o Log4j JMSSink
CVE-2022-23305	Não	O Amazon EMR não configura o Log4j JDBCAppender
CVE-2022-23307	Não	O Amazon EMR não configura o Log4j Chainsaw
CVE-2020-9493	Não	O Amazon EMR não configura o Log4j Chainsaw
CVE-2021-44832	Não	O Amazon EMR não configura o Log4j JDBCAppender com uma cadeia de conexão JNDI
CVE-2021-4104	Não	O Amazon EMR não usa o Log4j JMSAppender

CVE	Afeta o EMR	Observações
CVE-2020-9488	Não	Os aplicativos instalados pelo Amazon EMR não usam o Log4j SMTPAppender
CVE-2019-17571	Não	O Amazon EMR bloqueia o acesso público aos clusters e não inicia SocketServer
CVE-2019-17531	Não	Recomendamos que você faça a atualização para a versão mais recente do Amazon EMR. As versões 5.33.0 e posteriores do Amazon EMR usam o jackson-databind 2.6.7.4 ou posterior e as versões 6.1.0 e posteriores do EMR usam o jackson-databind 2.10.0 ou posterior. Essas versões do jackson-databind não são afetadas pelo CVE.

Arquivo de notas de versão do Amazon EMR

Notas de versão para todas as versões do Amazon EMR estão disponíveis a seguir. Para obter informações de versão abrangentes para cada versão, consulte Versões de lançamento 6.x do Amazon EMR, Versões de lançamento 5.x do Amazon EMR e Versões de lançamento 4.x do Amazon EMR.

Para receber atualizações quando uma nova versão do Amazon EMR estiver disponível, assine o RSS feed das notas de versão do Amazon EMR.

Versão 6.14.0

As notas da versão a seguir incluem informações para a versão 6.14.0 do Amazon EMR. As alterações são referentes à versão 6.13.0. Para obter informações sobre o cronograma da versão, consulte o Log de alterações 6.14.0.

Novos recursos

- Amazon EMR 6.14.0 supports Apache Spark 3.4.1, Apache Spark RAPIDS 23.06.0-amzn-2, Flink 1.17.1, Iceberg 1.3.1, and Trino 422.
- O Ajuste de Escala Gerenciado do Amazon EMR agora está disponível na região `ap-southeast-3` Ásia-Pacífico (Jacarta) para clusters criados com o Amazon EMR 6.14.0 e superior.

Alterações, melhorias e problemas resolvidos

- A versão 6.14.0 otimiza o gerenciamento de logs com o Amazon EMR em execução na Amazon. EC2 Como resultado, é possível ver uma pequena redução nos custos de armazenamento dos logs do cluster.
- A versão 6.14.0 melhora o fluxo de trabalho de ajuste de escala para considerar diferentes instâncias principais que têm variação substancial no tamanho dos volumes do Amazon EBS. Essa melhoria se aplica somente aos nós centrais; as operações de redução dos nós de tarefas não são afetadas.
- A versão 6.14.0 melhora a forma como o Amazon EMR interage com aplicações de código aberto, como. Apache Hadoop YARN ResourceManager and HDFS NameNode Essa melhoria reduz o risco de atrasos operacionais com o escalonamento do cluster e atenua as falhas de inicialização que ocorrem devido a problemas de conectividade com os aplicações de código aberto.
- A versão 6.14.0 otimiza a instalação da aplicação na inicialização do cluster. Isso melhora os tempos de inicialização do cluster para determinadas combinações de aplicações do Amazon EMR.
- A versão 6.14.0 corrige um problema em que as operações de redução de escala do cluster podem parar quando um cluster executado em uma VPC com um domínio personalizado é reiniciado no nó central ou da tarefa.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Malásia), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Malásia), Europa (Londres), Ásia-Pacífico (Melbourne), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (AWS GovCloud AWS GovCloud Oregon), México (Central) , Oeste dos EUA (Norte da

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Califórnia), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv), Ásia-Pacífico (Taipei), Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Hyderabad), Ásia-Pacífico (Jacarta) Arta), Ásia-Pacífico (Malásia), Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Malásia) , Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (EAU), Canadá (Central) , Oeste do Canadá

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Calgary), (Leste dos EUA), AWS GovCloud (Oeste dos EUA), China (AWS GovCloud Pequim), China (Ningxia), Israel (Tel Aviv), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Centro Leste (Emirados Árabes Unidos), Ásia-Pacífico (Melbourne),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Israel (Tel Aviv), Oeste do Canadá (Calgary), Europa (Zurique), Ásia-Pacífico (Malásia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 223.0	4.14.336	8 de março de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 131.0	4.14.336	14 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 124.0	4.14.336	7 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 109.0	4.14.334	24 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 218.0	4.14.330	2 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 206.0	4.14.330	22 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 116.0	4.14.328	11 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 101.0	4.14.327	17 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 906.0	4.14.322	11 de setembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

Versão 6.13.0

As notas da versão a seguir incluem informações para a versão 6.13.0 do Amazon EMR. As alterações são referentes à versão 6.12.0. Para obter informações sobre o cronograma da versão, consulte o Log de alterações 6.13.0.

Novos recursos

- Amazon EMR 6.13.0 supports Apache Spark 3.4.1, Apache Spark RAPIDS 23.06.0-amzn-1, CUDA Toolkit 11.8.0, and JupyterHub 1.5.0.

Alterações, melhorias e problemas resolvidos

- A versão 6.13.0 aprimora o daemon de gerenciamento de logs do Amazon EMR para garantir que todos os logs sejam carregados regularmente para o Amazon S3 quando um comando de encerramento de cluster é emitido. Isso facilita o encerramento mais rápido do cluster.
- A versão 6.13.0 aprimora os recursos de gerenciamento de logs do Amazon EMR para garantir o carregamento consistente e oportuno de todos os arquivos de log para o Amazon S3. Isso beneficia especialmente os clusters do EMR de execução prolongada.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Europa (Londres) , Ásia-Pacífico (Melbourn e), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (Oregon) AWS GovCloud , Oeste dos EUA (Norte da Califórnia), Ásia-Pací fico (AWS GovCloud Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv), Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha) , Europa (Estocolm o), Europa (Zurique) , Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Ásia-Pací fico (Jacarta) Ásia-Pacífico (Melbourne), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central), Oeste do Canadá (Calgary) AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud , (Leste dos EUA) AWS GovCloud , (Oeste dos EUA), China (Pequim), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Centro Leste (Emirados Árabes Unidos), Ásia-Pacífico (Melbourne),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Israel (Tel Aviv), Oeste do Canadá (Calgary), Europa (Zurique)
2.0.2024 001.0	4.14.352	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 816.0	4.14.350	21 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 809.0	4.14.349	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 719.0	4.14.348	25 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 223.0	4.14.336	8 de março de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 131.0	4.14.336	14 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 124.0	4.14.336	7 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 109.0	4.14.334	24 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 218.0	4.14.330	2 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 206.0	4.14.330	22 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 116.0	4.14.328	11 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 101.0	4.14.327	16 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 020.1	4.14.326	7 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 012.1	4.14.326	26 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 926.0	4.14.322	19 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 906.0	4.14.322	4 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 808.0	4.14.320	24 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

Versão 6.12.0

As notas da versão a seguir incluem informações para a versão 6.12.0 do Amazon EMR. As alterações são referentes à versão 6.11.0. Para obter informações sobre o cronograma da versão, consulte o Log de alterações 6.12.0.

Novos recursos

- Amazon EMR 6.12.0 supports Apache Spark 3.4.0, Apache Spark RAPIDS 23.06.0-amzn-0, CUDA 11.8.0, Apache Hudi 0.13.1-amzn-0, Apache Iceberg 1.3.0-amzn-0, Trino 414, and PrestoDB 0.281.
- As versões 6.12.0 e superiores do Amazon EMR oferecem suporte à integração LDAP com Apache Livy, Apache Hive through HiveServer 2 (), Trino, Presto e Hue. HS2 Você também pode instalar o Apache Spark e o Apache Hadoop em um cluster do EMR que use a versão 6.12.0 ou superior e configurá-los para usar LDAP. Para obter mais informações, consulte Usar servidores do Active Directory ou LDAP para autenticação com o Amazon EMR.

Alterações, melhorias e problemas resolvidos

- As versões 6.12.0 e superiores do Amazon EMR oferecem suporte ao runtime do Java 11 para o Flink. Para obter mais informações, consulte Configurar o Flink para ser executado com o Java 11.
- A versão 6.12.0 adiciona um mecanismo de nova tentativa ao fluxo de trabalho de escalabilidade de clusters para os clusters do EMR que executam Presto ou Trino. Essa melhoria reduz o risco de que o redimensionamento do cluster fique paralisado indefinidamente devido a uma única falha na operação de redimensionamento. Ela também aprimora a utilização dos clusters, porque seu cluster aumenta e reduz a escala verticalmente com mais rapidez.
- A versão 6.12.0 corrige um problema em que as operações de redução da escala verticalmente do cluster podem ficar paralisadas quando um nó central que está passando por uma desativação tranquila se torna não íntegro por qualquer motivo antes de ser totalmente desativado.
- A versão 6.12.0 melhora a lógica de redução da escala verticalmente do cluster para que o cluster não tente reduzir a escala verticalmente dos nós centrais abaixo da configuração do fator de replicação do HDFS para o cluster. Isso se alinha aos seus requisitos de redundância de dados e reduz a probabilidade de uma operação de escalabilidade paralisar.
- A versão 6.12.0 aprimora a performance e a eficiência do serviço de monitoramento de integridade do Amazon EMR, ao aumentar a velocidade com que ele registra em log as mudanças de estado das instâncias. Essa melhoria reduz a probabilidade de degradação do desempenho dos

nós do cluster que estão executando várias ferramentas de cliente ou aplicações de terceiros personalizadas.

- A versão 6.12.0 melhora a performance do daemon de gerenciamento de logs no cluster para o Amazon EMR. Como resultado, existe uma menor probabilidade de degradação da performance com clusters do EMR que executam etapas com alta simultaneidade.
- Com a versão 6.12.0 do Amazon EMR, o daemon de gerenciamento de logs foi atualizado para identificar todos os logs que estão em uso ativo com identificadores de arquivos abertos no armazenamento da instância local e nos processos associados. Essa atualização garante que o Amazon EMR exclua adequadamente os arquivos e recupere o espaço de armazenamento depois que os logs são arquivados no Amazon S3.
- A versão 6.12.0 inclui um aprimoramento do daemon de gerenciamento de logs que exclui diretórios de etapas vazios e não utilizados no sistema de arquivos de cluster local. Um número excessivamente grande de diretórios vazios pode degradar a performance dos daemons do Amazon EMR e resultar na utilização excessiva do disco.
- A versão 6.12.0 permite a alternância de logs do servidor de linha do tempo do YARN. Isso minimiza os cenários de utilização excessiva do disco, especialmente para clusters de execução prolongada.
- O tamanho padrão do volume raiz aumentou para 15 GB nas versões 6.10.0 e superiores do Amazon EMR. O tamanho padrão do volume raiz das versões anteriores é de 10 GB.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Europa (Londres) , Ásia-Pacífico (Melbourn e), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (Oregon) AWS GovCloud , Oeste dos EUA (Norte da Califórnia), Ásia-Pací fico (AWS GovCloud Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv), Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha) , Europa (Estocolm o), Europa (Zurique) , Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Ásia-Pací fico (Jacarta) Ásia-Pacífico (Melbourne), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central), Oeste do Canadá (Calgary) AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud , (Leste dos EUA) AWS GovCloud , (Oeste dos EUA), China (Pequim), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Centro Leste (Emirados Árabes Unidos), Ásia-Pacífico (Melbourne),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Israel (Tel Aviv), Oeste do Canadá (Calgary), Europa (Zurique)
2.0.2024 001.0	4.14.352	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 816.0	4.14.350	21 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 809.0	4.14.349	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 719.0	4.14.348	25 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 223.0	4.14.336	8 de março de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 131.0	4.14.336	14 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 124.0	4.14.336	7 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 109.0	4.14.334	24 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 218.0	4.14.330	2 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 206.0	4.14.330	22 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 116.0	4.14.328	11 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 101.0	4.14.327	16 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 020.1	4.14.326	7 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 012.1	4.14.326	26 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 926.0	4.14.322	19 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 906.0	4.14.322	4 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 822.0	4.14.322	30 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 808.0	4.14.320	24 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 727.0	4.14.320	14 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 719.0	4.14.320	2 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 628.0	4.14.318	12 de julho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta) , África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central)

Versão 6.11.1

As notas da versão a seguir incluem informações para a versão 6.11.1 do Amazon EMR. As alterações são referentes à versão 6.11.0. Para obter informações sobre o cronograma da versão, consulte o Log de alterações 6.11.1.

Alterações, melhorias e problemas resolvidos

- Devido à contenção de bloqueio, um nó pode entrar em um deadlock se for adicionado ou removido ao mesmo tempo em que tenta ser desativado. Como resultado, o Hadoop Resource Manager (YARN) deixa de responder e afeta todos os contêineres de entrada e em execução no momento.
- Esta versão inclui uma alteração que permite que clusters de alta disponibilidade se recuperem de um estado de falha após a reinicialização.
- Esta versão inclui correções de segurança para Hue e. HBase
- Esta versão corrige um problema em que clusters que estão executando workloads no Spark com o Amazon EMR podem receber silenciosamente resultados incorretos com `contains`, `startsWith`, `endsWith` e `like`. Esse problema ocorre quando você usa as expressões em campos particionados que têm metadados no Hive3 Metastore Server (HMS) no Amazon EMR.
- Esta versão corrige um problema com controle de utilização no lado do Glue quando não há funções definidas pelo usuário (UDF).
- Esta versão corrige um problema que exclui logs de contêineres pelo serviço de agregação de logs de nó antes que o pusher de logs possa enviá-los para o S3 em caso de desativação do YARN.
- Esta versão corrige um problema com as métricas do FairShare Scheduler quando o Node Label está habilitado para o Hadoop.
- Esta versão corrige um problema que afetou a performance do Spark quando você definiu um valor de `true` padrão para a configuração `spark.yarn.heterogeneousExecutors.enabled` no `spark-defaults.conf`.
- Esta versão corrige um problema com a falha do Reduce Task em ler dados embaralhados. O problema causou falhas na consulta do Hive com um erro de memória corrompida.
- Esta versão adiciona um mecanismo de nova tentativa ao fluxo de trabalho de escalabilidade de clusters para clusters do EMR que executam o Presto ou o Trino. Essa melhoria reduz o risco de que o redimensionamento do cluster fique paralisado indefinidamente devido a uma única falha na operação de redimensionamento. Ela também aprimora a utilização dos clusters, porque seu cluster aumenta e reduz a escala verticalmente com mais rapidez.

- Esta versão melhora a lógica de redução da escala verticalmente do cluster para que o cluster não tente reduzir a escala verticalmente dos nós centrais abaixo da configuração do fator de replicação do HDFS para o cluster. Isso se alinha aos seus requisitos de redundância de dados e reduz a probabilidade de uma operação de escalabilidade paralisar.
- O daemon de gerenciamento de logs foi atualizado para identificar todos os logs que estão em uso ativo com identificadores de arquivos abertos no armazenamento da instância local e nos processos associados. Essa atualização garante que o Amazon EMR exclua adequadamente os arquivos e recupere o espaço de armazenamento depois que os logs são arquivados no Amazon S3.
- Esta versão inclui um aprimoramento do daemon de gerenciamento de logs que exclui diretórios de etapas vazios e não utilizados no sistema de arquivos de cluster local. Um número excessivamente grande de diretórios vazios pode degradar a performance dos daemons do Amazon EMR e resultar na utilização excessiva do disco.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Malásia), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Malásia), Europa (Londres), Ásia-Pacífico (Melbourne), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (AWS GovCloud AWS GovCloud Oregon), México (Central) , Oeste dos EUA (Norte da Califórnia), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv), Ásia-Pacífico (Taipei), Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Hyderabad), Ásia-Pacífico (Jacarta) Arta), Ásia-Pacífico (Malásia), Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Malásia) , Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (EAU), Canadá (Central) , Oeste do Canadá

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Calgary), (Leste dos EUA), AWS GovCloud (Oeste dos EUA), China (AWS GovCloud Pequim), China (Ningxia), Israel (Tel Aviv), México (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Centro Leste (Emirados Árabes Unidos), Ásia-Pacífico (Melbourne),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Israel (Tel Aviv), Oeste do Canadá (Calgary), Europa (Zurique), Ásia-Pacífico (Malásia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 001.0	4.14.352	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 816.0	4.14.350	21 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 809.0	4.14.349	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 719.0	4.14.348	25 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 223.0	4.14.336	8 de março de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 131.0	4.14.336	14 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 124.0	4.14.336	7 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 109.0	4.14.334	24 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia) , Oeste do Canadá (Calgary)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 218.0	4.14.330	2 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 206.0	4.14.330	22 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 116.0	4.14.328	11 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 101.0	4.14.327	16 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 020.1	4.14.326	7 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 012.1	4.14.326	26 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 926.0	4.14.322	19 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderaba d) Abad), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pací fico (Osaka), Ásia-Pací fico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv), AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud (Oeste dos EUA), AWS GovCloud (Leste dos EUA), China (Pequim), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 906.0	4.14.322	4 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 822.0	4.14.322	30 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 808.0	4.14.320	24 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 727.0	4.14.320	14 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central)

Versão 6.11.0

As notas da versão a seguir incluem informações para a versão 6.11.0 do Amazon EMR. As alterações são referentes à versão 6.10.0. Para obter informações sobre o cronograma da versão, consulte o log de alterações.

Novos recursos

- O Amazon EMR 6.11.0 é compatível com Apache Spark 3.3.2-amzn-0, Apache Spark RAPIDS 23.02.0-amzn-0, CUDA 11.8.0, Apache Hudi 0.13.0-amzn-0, Apache Iceberg 1.2.0-amzn-0, Trino 410-amzn-0 e PrestoDB 0.279-amzn-0.

Alterações, melhorias e problemas resolvidos

- Com o Amazon EMR 6.11.0, o conector do DynamoDB foi atualizado para a versão 5.0.0. A versão 5.0.0 usa AWS SDK for Java 2.x. As versões anteriores usavam AWS SDK para Java 1.x. Como resultado dessa atualização, recomendamos firmemente que você teste seu código antes de usar o conector do DynamoDB com o Amazon EMR 6.11.
- Quando o conector do DynamoDB para Amazon EMR 6.11.0 chama o serviço do DynamoDB, ele usa o valor da região que você fornece para a propriedade `dynamodb.endpoint`. Recomendamos que você também configure `dynamodb.region` quando usar `dynamodb.endpoint` e que ambas as propriedades tenham como destino a mesma Região da AWS. Se você usar `dynamodb.endpoint` e não configurar`dynamodb.region`, o conector do DynamoDB para Amazon EMR 6.11.0 retornará uma exceção de região inválida e tentará reconciliar suas Região da AWS informações do serviço de metadados de instância da Amazon (IMDS). EC2 Se o conector não conseguir recuperar a região do IMDS, o padrão será Leste dos EUA (Norte da Virgínia) (`us-east-1`). O erro a seguir é um exemplo da exceção de região inválida que você pode obter se não configurar adequadamente a `dynamodb.region` propriedade: `error software.amazon.awssdk.services.dynamodb.model.DynamoDbException: Credential should be scoped to a valid region.` Para obter mais informações sobre as classes afetadas pela AWS SDK para Java atualização para 2.x, consulte o commit Upgrade AWS SDK para Java from 1.x to 2.x (#175) no GitHub repositório do conector Amazon EMR - DynamoDB.
- Esta versão corrige um problema em que os dados da coluna se tornam `NULL` quando você usa o Delta Lake para armazenar dados da tabela Delta no Amazon S3 após a operação de renomeação da coluna. Para obter mais informações sobre esse atributo experimental no Delta Lake, consulte Operação de renomeação de coluna no Guia do usuário do Delta Lake.
- A versão 6.11.0 corrige um problema que pode ocorrer quando você cria um nó de borda ao replicar um dos nós primários de um cluster com vários nós primários. O nó de borda replicado pode causar atrasos nas operações de redução da escala verticalmente ou resultar em alta utilização de memória nos nós primários. Para obter mais informações sobre como criar um nó de

borda para se comunicar com seu cluster EMR, consulte Edge Node Creator no `aws-samples` repositório em. GitHub

- A versão 6.11.0 melhora o processo de automação que o Amazon EMR usa para remontar volumes do Amazon EBS em uma instância após uma reinicialização.
- A versão 6.11.0 corrige um problema que resultou em lacunas intermitentes nas métricas do Hadoop que o Amazon EMR publica na Amazon. CloudWatch
- A versão 6.11.0 corrige um problema com clusters do EMR em que uma atualização no arquivo de configuração do YARN que contém a lista de exclusão de nós do cluster é interrompida devido à utilização excessiva do disco. A atualização incompleta impede futuras operações de redução da escala verticalmente do cluster. Esta versão garante que o cluster permaneça íntegro e que as operações de escalabilidade funcionem conforme esperado.
- O tamanho padrão do volume raiz aumentou para 15 GB nas versões 6.10.0 e superiores do Amazon EMR. O tamanho padrão do volume raiz das versões anteriores é de 10 GB.
- O Hadoop 3.3.3 introduziu uma alteração no YARN (YARN-9608) que mantém os nós em que os contêineres eram executados em um estado de desativação até que a aplicação seja concluída. Essa alteração garante que dados locais, como dados embaralhados, não sejam perdidos e que você não precise executar o trabalho novamente. Essa abordagem também pode levar à subutilização de recursos em clusters com ou sem o ajuste de escala gerenciado habilitado.

  Com as versões 6.11.0 e superiores do Amazon EMR, além das versões 6.8.1, 6.9.1 e 6.10.1, o valor de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` é definido como `false` em `yarn-site.xml` para resolver esse problema.

  Embora a correção resolva os problemas introduzidos pelo YARN-9608, ela pode fazer com que os trabalhos do Hive falhem devido à perda de dados embaralhados em clusters com ajuste de escala gerenciado habilitado. Reduzimos esse risco nesta versão também ao configurar `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data` para workloads do Hive. Essa configuração só está disponível com as versões 6.11.0 e superiores do Amazon EMR.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

> **Note**
>
> Essa versão não recebe mais atualizações automáticas da AMI, pois foi substituída por uma ou mais versões de patch. A versão de patch é indicada pelo número após o segundo ponto decimal (`6.8.1`). Para ver se você está usando a versão de patch mais recente, verifique as versões disponíveis no Guia de lançamento ou verifique o menu suspenso de versões do Amazon EMR quando criar um cluster no console ou use a ação de API `ListReleaseLabels` ou da CLI `list-release-labels`. Para obter atualizações sobre novas versões, assine o feed RSS na página Novidades.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Europa (Londres) , Ásia-Pacífico (Melbourn e), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (Oregon) AWS GovCloud , Oeste dos EUA (Norte da Califórnia), Ásia-Pací fico (AWS GovCloud Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha) , Europa (Estocolm o), Europa (Zurique) , Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Ásia-Pací fico (Jacarta) Ásia-Pacífico (Melbourne), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central), Oeste do Canadá (Calgary) AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud , (Leste dos EUA) AWS GovCloud , (Oeste dos EUA), China (Pequim), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Centro Leste (Emirados Árabes Unidos), Ásia-Pacífico (Melbourne),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Israel (Tel Aviv), Oeste do Canadá (Calgary), Europa (Zurique)
2.0.2024 001.0	4.14.352	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 816.0	4.14.350	21 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 809.0	4.14.349	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 719.0	4.14.348	25 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)
2.0.2023 808.0	4.14.320	24 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), Israel (Tel Aviv).

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 727.0	4.14.320	14 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 719.0	4.14.320	2 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 628.0	4.14.318	12 de julho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Zurique), Europa (Milão), Europa (Espanha) , Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 612.0	4.14.314	23 de junho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Zurique), Europa (Milão), Europa (Espanha) , Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 504.1	4.14.313	16 de maio de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta) , África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central)

Versão 6.10.0

As notas da versão a seguir incluem informações para a versão 6.10.0 do Amazon EMR. As alterações são referentes à versão 6.9.0. Para obter informações sobre o cronograma da versão, consulte o log de alterações.

Novos recursos

- O Amazon EMR 6.10.0 é compatível com Apache Spark 3.3.1, Apache Spark RAPIDS 22.12.0, CUDA 11.8.0, Apache Hudi 0.12.2-amzn-0, Apache Iceberg 1.1.0-amzn-0, Trino 403 e PrestoDB 0.278.1.
- O Amazon EMR 6.10.0 inclui um conector Trino-Hudi nativo que fornece acesso de leitura aos dados nas tabelas Hudi. Você pode ativar o conector com `trino-cli --catalog hudi` e configurar o conector de acordo com suas necessidades com `trino-connector-hudi`. A integração nativa com o Amazon EMR significa que você não precisa mais usar `trino-connector-hive` para consultar tabelas do Hudi. Para obter uma lista das configurações compatíveis com o novo conector, consulte a página do conector do Hudi na documentação do Trino.
- O Amazon EMR 6.10.0 e versões posteriores oferecem suporte à integração do Apache Zeppelin com o Apache Flink. Consulte Usar trabalhos do Flink pelo Zeppelin no Amazon EMR para obter mais informações.

Problemas conhecidos

- O Hadoop 3.3.3 introduziu uma alteração no YARN (YARN-9608) que mantém os nós em que os contêineres eram executados em um estado de desativação até que a aplicação seja concluída. Essa alteração garante que dados locais, como dados embaralhados, não sejam perdidos e que você não precise executar o trabalho novamente. Essa abordagem também pode levar à subutilização de recursos em clusters com ou sem o ajuste de escala gerenciado habilitado.

  Para contornar esse problema no Amazon EMR 6.10.0, você pode definir o valor de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` como `false` em `yarn-site.xml`. Nas versões 6.11.0 e superiores do Amazon EMR, além das versões 6.8.1, 6.9.1 e 6.10.1, o config é definido como `false` por padrão para resolver esse problema.

Alterações, melhorias e problemas resolvidos

- O Amazon EMR 6.10.0 remove a dependência de `minimal-json.jar` para a integração do Amazon Redshift para Apache Spark e adiciona automaticamente os jars necessários relacionados ao Spark-Redshift ao caminho de classe do executor para o Spark: `spark-redshift.jar`, `spark-avro.jar` e `RedshiftJDBC.jar`.
- A versão 6.10.0 aprimora o daemon de gerenciamento de logs no cluster para monitorar pastas de log adicionais no cluster do EMR. Essa melhoria minimiza os cenários de utilização excessiva do disco.
- A versão 6.10.0 reinicia automaticamente o daemon de gerenciamento de logs no cluster quando ele é interrompido. Essa melhoria reduz o risco de os nós parecerem não íntegros devido à utilização excessiva do disco.
- O Amazon EMR 6.10.0 é compatível com endpoints regionais para mapeamento de usuários do EMRFS.
- O tamanho padrão do volume raiz aumentou para 15 GB nas versões 6.10.0 e superiores do Amazon EMR. O tamanho padrão do volume raiz das versões anteriores é de 10 GB.
- A versão 6.10.0 corrige um problema que fazia com que os trabalhos do Spark paralisassem quando todos os executores restantes do Spark estivessem em um host em desativação com o gerenciador de recursos do YARN.
- Com o Amazon EMR 6.6.0 a 6.9.x, as consultas INSERT com partição dinâmica e uma cláusula ORDER BY ou SORT BY sempre terá dois redutores. Esse problema é causado pela alteração do OSS HIVE-20703, que coloca a otimização da partição dinâmica de classificação sob uma decisão baseada em custos. Se sua workload não exigir a classificação de partições dinâmicas, recomendamos que você defina a propriedade `hive.optimize.sort.dynamic.partition.threshold` como `-1` para desabilitar o novo atributo e obter o número de redutores calculado corretamente. Esse problema foi corrigido no OSS Hive como parte do HIVE-22269 e foi corrigido no Amazon EMR 6.10.0.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

> **Note**
>
> Essa versão não recebe mais atualizações automáticas da AMI, pois foi substituída por uma ou mais versões de patch. A versão de patch é indicada pelo número após o segundo

ponto decimal (`6.8.1`). Para ver se você está usando a versão de patch mais recente, verifique as versões disponíveis no Guia de lançamento ou verifique o menu suspenso de versões do Amazon EMR quando criar um cluster no console ou use a ação de API `ListReleaseLabels` ou da CLI `list-release-labels`. Para obter atualizações sobre novas versões, assine o feed RSS na página Novidades.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Europa (Londres) , Ásia-Pacífico (Melbourn e), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (Oregon) AWS GovCloud , Oeste dos EUA (Norte da Califórnia), Ásia-Pací fico (AWS GovCloud Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha) , Europa (Estocolm o), Europa (Zurique) , Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Ásia-Pací fico (Jacarta) Ásia-Pacífico (Melbourne), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central), Oeste do Canadá (Calgary) AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud , (Leste dos EUA) AWS GovCloud , (Oeste dos EUA), China (Pequim), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Centro Leste (Emirados Árabes Unidos), Ásia-Pacífico (Melbourne),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Israel (Tel Aviv), Oeste do Canadá (Calgary), Europa (Zurique)
2.0.2024 001.0	4.14.352	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 816.0	4.14.350	21 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 809.0	4.14.349	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 719.0	4.14.348	25 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)
2.0.2023 808.0	4.14.320	24 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), Israel (Tel Aviv).

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 727.0	4.14.320	14 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 719.0	4.14.320	2 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 628.0	4.14.318	12 de julho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Zurique), Europa (Milão), Europa (Espanha) , Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 612.0	4.14.314	23 de junho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Zurique), Europa (Milão), Europa (Espanha) , Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 504.1	4.14.313	16 de maio de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Zurique), Europa (Milão), Europa (Espanha) , Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 418.0	4.14.311	3 de maio de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Zurique), Europa (Milão), Europa (Espanha), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 404.1	4.14.311	18 de abril de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos)
2.0.2023 404.0	4.14.311	10 de abril de 2023	Leste dos EUA (Norte da Virgínia), Europa (Paris)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 320.0	4.14.309	30 de março de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 207.0	4.14.304	22 de fevereiro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos)

Versão 6.9.0

As notas da versão a seguir incluem informações para a versão 6.9.0 do Amazon EMR. As alterações são referentes à versão 6.8.0 do Amazon EMR. Para obter informações sobre o cronograma da versão, consulte o log de alterações.

Novos atributos

- A versão 6.9.0 do Amazon EMR é compatível com Apache Spark RAPIDS 22.08.0, Apache Hudi 0.12.1, Apache Iceberg 0.14.1, Trino 398 e Tez 0.10.2.
- A versão 6.9.0 do Amazon EMR inclui uma nova aplicação de código aberto, Delta Lake 2.1.0.
- A integração do Amazon Redshift para Apache Spark está inclusa nas versões 6.9.0 e posteriores do Amazon EMR. Anteriormente uma ferramenta de código aberto, a integração nativa é um conector do Spark que você pode usar para criar aplicações do Apache Spark que realizam a leitura e a gravação de dados no Amazon Redshift e no Amazon Redshift sem servidor. Para obter mais informações, consulte Usar a integração do Amazon Redshift para Apache Spark com o Amazon EMR .
- A versão 6.9.0 do Amazon EMR adiciona suporte ao arquivamento de logs no Amazon S3 durante a redução da escala do cluster verticalmente. Anteriormente, só era possível arquivar arquivos de log no Amazon S3 durante o encerramento do cluster. A nova capacidade garante que os arquivos de log gerados no cluster persistam no Amazon S3 mesmo após o encerramento do nó. Para obter mais informações, consulte Configurar registro em log e depuração do cluster.
- Para dar suporte a consultas de longa execução, o Trino agora inclui um mecanismo de execução tolerante a falhas. A execução tolerante a falhas atenua as falhas nas consultas ao tentar novamente as consultas com falha ou as tarefas dos seus componentes.
- Você pode usar o Apache Flink no Amazon EMR para `BATCH` unificado e processamento de `STREAM` de tabelas do Apache Hive ou metadados de qualquer tablesource do Flink, como Iceberg, Kinesis ou Kafka. Você pode especificar o AWS Glue Data Catalog como metastore para o Flink usando a API AWS Management Console,, AWS CLI ou Amazon EMR. Para obter mais informações, consulte Configurar o Flink no Amazon EMR.
- Agora você pode especificar funções de tempo de execução AWS Identity and Access Management (IAM) e controle de acesso AWS Lake Formation baseado para consultas do Apache Spark, Apache Hive e Presto no Amazon EC2 EMR em clusters com o Amazon AI Studio. SageMaker Para obter mais informações, consulte Configurar funções de runtime para as etapas do Amazon EMR.

Problemas conhecidos

- Para a versão 6.9.0 do Amazon EMR, o Trino não funciona em clusters habilitados para o Apache Ranger. Se você precisar usar o Trino com o Ranger, entre em contato com o Suporte.

- Se você usar a integração do Amazon Redshift para Apache Spark e tiver um time, timetz, timestamp ou timestamptz com precisão de microssegundos no formato Parquet, o conector arredondará os valores de tempo para o valor de milissegundo mais próximo. Como solução alternativa, use o parâmetro `unload_s3_format` do formato de descarregamento de texto.
- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
  - O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

  Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.
- As conexões com clusters do Amazon EMR do Amazon SageMaker AI Studio podem falhar intermitentemente com um código de resposta 403 Forbidden. Esse erro ocorre quando a configuração do perfil do IAM no cluster leva mais de 60 segundos. Como solução alternativa, você pode instalar um patch do Amazon EMR para permitir novas tentativas e aumentar o tempo limite para um mínimo de 300 segundos. Use as etapas a seguir para aplicar a ação de bootstrap quando iniciar o cluster.

  1. Baixe o script de bootstrap e os arquivos RPM do Amazon URIs S3 a seguir.

     ```
     s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/gcsc/replace-rpms.sh
     s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/gcsc/emr-secret-agent-1.18.0-SNAPSHOT20221121212949.noarch.rpm
     ```

  2. Carregue os arquivos da etapa anterior em um bucket do Amazon S3 de sua propriedade. O bucket deve estar no mesmo Região da AWS local em que você planeja iniciar o cluster.

3. Inclua a seguinte ação de bootstrap ao iniciar o cluster do EMR. Substitua *bootstrap_URI* e *RPM_URI* pelo correspondente URIs do Amazon S3.

```
--bootstrap-actions "Path=bootstrap_URI,Args=[RPM_URI]"
```

- Com as versões 5.36.0 e 6.6.0 a 6.9.0 do Amazon EMR, os componentes do serviço `SecretAgent` e `RecordServer` podem sofrer perda de dados de log devido a uma configuração incorreta do padrão de nome de arquivo nas propriedades do Log4j2. A configuração incorreta faz com que os componentes gerem somente um arquivo de log por dia. Quando a estratégia de rotação ocorre, ela substitui o arquivo existente em vez de gerar um novo arquivo de log, conforme esperado. Como solução alternativa, use uma ação de bootstrap para gerar arquivos de log a cada hora e acrescentar um número inteiro de incremento automático no nome do arquivo para lidar com a rotação.

  Para as versões 6.6.0 a 6.9.0 do Amazon EMR, use a seguinte ação de bootstrap ao iniciar um cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

  Para o Amazon EMR 5.36.0, use a ação de bootstrap a seguir ao iniciar um cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

- O Apache Flink fornece FileSystem conectores nativos S3 FileSystem e Hadoop, que permitem que os aplicativos criem FileSink e gravem os dados no Amazon S3. Isso FileSink falha com uma das duas exceções a seguir.

```
java.lang.UnsupportedOperationException: Recoverable writers on Hadoop are only
 supported for HDFS
```

```
Caused by: java.lang.NoSuchMethodError:
 org.apache.hadoop.io.retry.RetryPolicies.retryOtherThanRemoteAndSaslException(Lorg/
apache/hadoop/io/retry/RetryPolicy;Ljava/util/Map;)Lorg/apache/hadoop/io/retry/
RetryPolicy;
```

```
                                        at
 org.apache.hadoop.yarn.client.RMProxy.createRetryPolicy(RMProxy.java:302) ~[hadoop-
yarn-common-3.3.3-amzn-0.jar:?]
```

Como solução alternativa, você pode instalar um patch do Amazon EMR, que corrige o problema acima no Flink. Para aplicar a ação de bootstrap quando iniciar o cluster, execute as etapas a seguir.

1. Baixe o flink-rpm no bucket Amazon S3. Seu caminho de RPM é `s3://DOC-EXAMPLE-BUCKET/rpms/flink/`.
2. Baixe o script de bootstrap e os arquivos RPM do Amazon S3 usando o URI a seguir. *regionName*Substitua pelo Região da AWS local em que você planeja iniciar o cluster.

   ```
   s3://emr-data-access-control-regionName/customer-bootstrap-actions/gcsc/replace-rpms.sh
   ```
3. O Hadoop 3.3.3 introduziu uma alteração no YARN (YARN-9608) que mantém os nós em que os contêineres eram executados em um estado de desativação até que a aplicação seja concluída. Essa alteração garante que dados locais, como dados embaralhados, não sejam perdidos e que você não precise executar o trabalho novamente. Nas versões 6.8.0 e 6.9.0 do Amazon EMR, essa abordagem também pode levar à subutilização de recursos em clusters com ou sem o ajuste de escala gerenciado habilitado.

   Com o Amazon EMR 6.10.0, há uma solução alternativa para esse problema: definir o valor de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` como `false` em `yarn-site.xml`. Nas versões 6.11.0 e superiores do Amazon EMR, além das versões 6.8.1, 6.9.1 e 6.10.1, o config é definido como `false` por padrão para resolver esse problema.

Alterações, melhorias e problemas resolvidos

- Para as versões 6.9.0 e posteriores do Amazon EMR, todos os componentes instalados pelo Amazon EMR que usam bibliotecas do Log4j usam o Log4j versão 2.17.1 ou posterior.
- Ao usar o conector DynamoDB com o Spark nas versões 6.6.0, 6.7.0 e 6.8.0 do Amazon EMR, todas as leituras da tabela retornam um resultado vazio, mesmo que a divisão de entrada faça referência a dados que não estão vazios. A versão 6.9.0 do Amazon EMR corrige esse problema.
- O Amazon EMR 6.9.0 adiciona suporte limitado ao controle de acesso baseado no Lake Formation com o Apache Hudi ao ler dados usando o Spark SQL. O suporte se destina a consultas SELECT

usando o Spark SQL e é limitado ao controle de acesso em nível de coluna. Para obter mais informações, consulte Hudi e Lake Formation.

- Quando você usa o Amazon EMR 6.9.0 para criar um cluster do Hadoop com Rótulos de nós habilitados, a API de métricas do YARN retorna informações agregadas em todas as partições, em vez de na partição padrão. Para obter mais informações, consulte YARN-11414.
- Com a versão 6.9.0 do Amazon EMR, atualizamos o Trino para a versão 398, que usa Java 17. A versão anterior do Trino compatível com o Amazon EMR 6.8.0 era Trino 388 em execução no Java 11. Para obter mais informações sobre essa alteração, consulte Atualizações do Trino para Java 17 no blog do Trino.
- Esta versão corrige um problema de incompatibilidade de sequência de tempo entre o Apache BigTop e o Amazon EMR na EC2 sequência de inicialização do cluster. Essa incompatibilidade de sequência de tempo ocorre quando um sistema tenta realizar duas ou mais operações ao mesmo tempo em vez de fazê-las na sequência correta. Como resultado, determinadas configurações de cluster apresentaram tempos limite de inicialização da instância e tempos de inicialização do cluster mais lentos.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

> **Note**
>
> Essa versão não recebe mais atualizações automáticas da AMI, pois foi substituída por uma ou mais versões de patch. A versão de patch é indicada pelo número após o segundo ponto decimal (`6.8.1`). Para ver se você está usando a versão de patch mais recente, verifique as versões disponíveis no Guia de lançamento ou verifique o menu suspenso de versões do Amazon EMR quando criar um cluster no console ou use a ação de API `ListReleaseLabels` ou da CLI `list-release-labels`. Para obter atualizações sobre novas versões, assine o feed RSS na página Novidades.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Europa (Londres) , Ásia-Pacífico (Melbourn e), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (Oregon) AWS GovCloud , Oeste dos EUA (Norte da Califórnia), Ásia-Pací fico (AWS GovCloud Cingapura), Ásia-Pacífico (Sydney), Israel (Tel Aviv),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Canadá (Central), Oeste do Canadá (Calgary) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha) , Europa (Estocolm o), Europa (Zurique) , Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderabad), Ásia-Pací fico (Jacarta) Ásia-Pacífico (Melbourne), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central), Oeste do Canadá (Calgary) AWS

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			GovCloud , (Leste dos EUA) AWS GovCloud , (Oeste dos EUA), China (Pequim), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Centro Leste (EAU), Ásia-Pacífico (Melbourne),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Israel (Tel Aviv), Europa (Zurique)
2.0.2024 001.0	4.14.352	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 816.0	4.14.350	21 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 809.0	4.14.349	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 719.0	4.14.348	25 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv), Oeste do Canadá (Calgary)
2.0.2023 808.0	4.14.320	24 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourne), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), Israel (Tel Aviv).

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 727.0	4.14.320	14 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 719.0	4.14.320	2 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , Israel (Tel Aviv)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 628.0	4.14.318	12 de julho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 612.0	4.14.314	23 de junho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 504.1	4.14.313	16 de maio de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 418.0	4.14.311	3 de maio de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 404.1	4.14.311	18 de abril de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)
2.0.2023 404.0	4.14.311	10 de abril de 2023	Leste dos EUA (Norte da Virgínia), Europa (Paris)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 320.0	4.14.309	30 de março de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 307.0	4.14.305	15 de março de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 207.0	4.14.304	22 de fevereiro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 210.1	4.14.301	12 de janeiro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 103.3	4.14.296	5 de dezembro de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

Versão 6.8.0

As notas da versão a seguir incluem informações para a versão 6.8.0 do Amazon EMR. As alterações são referentes à versão 6.7.0.

Novos atributos

- O recurso de etapas do Amazon EMR agora oferece suporte a endpoints e clientes Apache Livy. JDBC/ODBC Para obter mais informações, consulte Configurar funções de runtime para as etapas do Amazon EMR.
- A versão 6.8.0 do Amazon EMR vem com a versão 2.4.12 do Apache. HBase Com essa HBase versão, você pode arquivar e excluir suas HBase tabelas. O processamento de arquivos do Amazon S3 renomeia todos os arquivos da tabela para o diretório de arquivos. Isso pode ser um processo custoso e demorado. Agora, você pode pular o processamento de arquivos e rapidamente eliminar e excluir tabelas grandes. Para obter mais informações, consulte Usando a HBase concha.

Problemas conhecidos

- O Hadoop 3.3.3 introduziu uma alteração no YARN (YARN-9608) que mantém os nós em que os contêineres eram executados em um estado de desativação até que a aplicação seja concluída. Essa alteração garante que dados locais, como dados embaralhados, não sejam perdidos e que você não precise executar o trabalho novamente. Nas versões 6.8.0 e 6.9.0 do Amazon EMR, essa abordagem também pode levar à subutilização de recursos em clusters com ou sem o ajuste de escala gerenciado habilitado.

  Com o Amazon EMR 6.10.0, há uma solução alternativa para esse problema: definir o valor de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` como `false` em `yarn-site.xml`. Nas versões 6.11.0 e superiores do Amazon EMR, além das versões 6.8.1, 6.9.1 e 6.10.1, o config é definido como `false` por padrão para resolver esse problema.

Alterações, melhorias e problemas resolvidos

- Quando a versão 6.5.0, 6.6.0 ou 6.7.0 do Amazon EMR leu as tabelas do Apache Phoenix por meio do shell do Apache Spark, o Amazon EMR produziu um `NoSuchMethodError`. A versão 6.8.0 do Amazon EMR corrige esse problema.
- A versão 6.8.0 do Amazon EMR vem com o Apache Hudi 0.11.1; no entanto, os clusters do Amazon EMR 6.8.0 também são compatíveis com o código aberto `hudi-spark3.3-bundle_2.12` do Hudi 0.12.0.

- A versão 6.8.0 do Amazon EMR vem com a versão 3.3.0 do Apache Spark. Esta versão do Spark usa o Apache Log4j 2 e o arquivo `log4j2.properties` para configurar o Log4j nos processos do Spark. Se você usar o Spark no cluster ou criar clusters do EMR com parâmetros de configuração personalizados e quiser atualizar para a versão 6.8.0 do Amazon EMR, deverá migrar para a nova classificação de configuração `spark-log4j2` e para o formato de chave do Apache Log4j 2. Para obter mais informações, consulte Migrar do Apache Log4j 1.x para Log4j 2.x.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

> **Note**
>
> Essa versão não recebe mais atualizações automáticas da AMI, pois foi substituída por uma ou mais versões de patch. A versão de patch é indicada pelo número após o segundo ponto decimal (`6.8.1`). Para ver se você está usando a versão de patch mais recente, verifique as versões disponíveis no Guia de lançamento ou verifique o menu suspenso de versões do Amazon EMR quando criar um cluster no console ou use a ação de API `ListReleaseLabels` ou da CLI `list-release-labels`. Para obter atualizações sobre novas versões, assine o feed RSS na página Novidades.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Europa (Londres) , Ásia-Pacífico (Melbourn e), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (Oregon) AWS GovCloud , Oeste dos EUA (Norte da Califórnia), Ásia-Pací fico (AWS GovCloud Cingapura), Ásia-Pací fico (Sydney), Israel (Tel Aviv), Canadá (Central) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Hyderabad), Ásia-Pacífico (Jacarta) Ásia-Pacífico (Melbourne), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central) , AWS GovCloud (Leste dos EUA), AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Oeste dos EUA), China (Pequim), China (Ningxia), Israel (Tel Aviv)
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Centro Leste (EAU), Ásia-Pacífico (Melbourne),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Israel (Tel Aviv), Europa (Zurique)
2.0.2024 001.0	4.14.352	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 816.0	4.14.350	21 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 809.0	4.14.349	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 719.0	4.14.348	25 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU), Europa (Espanha), Europa

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Zurique), Ásia-Pacífico (Melbourne), Israel (Tel Aviv)
2.0.2023 808.0	4.14.320	24 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórni a), Oeste dos EUA (Oregon), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 727.0	4.14.320	14 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 719.0	4.14.320	2 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), Ásia-Pacífico (Melbourn e), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 628.0	4.14.318	12 de julho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 612.0	4.14.314	23 de junho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 504.1	4.14.313	16 de maio de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 418.0	4.14.311	3 de maio de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 404.1	4.14.311	18 de abril de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)
2.0.2023 404.0	4.14.311	10 de abril de 2023	Leste dos EUA (Norte da Virgínia), Europa (Paris)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 320.0	4.14.309	30 de março de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 307.0	4.14.305	15 de março de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 207.0	4.14.304	22 de fevereiro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 119.1	4.14.301	3 de fevereiro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 210.1	4.14.301	22 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 103.3	4.14.296	5 de dezembro de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 004.0	4.14.294	2 de novembro de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 912.1	4.14.291	6 de setembro de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

Problemas conhecidos

- Ao usar o conector DynamoDB com o Spark nas versões 6.6.0, 6.7.0 e 6.8.0 do Amazon EMR, todas as leituras da tabela retornam um resultado vazio, mesmo que a divisão de entrada faça referência a dados que não estão vazios. Isso ocorre porque o Spark 3.2.0 define `spark.hadoopRDD.ignoreEmptySplits` como `true` por padrão. Como solução alternativa,

defina explicitamente `spark.hadoopRDD.ignoreEmptySplits` como `false`. A versão 6.9.0 do Amazon EMR corrige esse problema.

- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
  - O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

  Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.
- Com as versões 5.36.0 e 6.6.0 a 6.9.0 do Amazon EMR, os componentes do serviço `SecretAgent` e `RecordServer` podem sofrer perda de dados de log devido a uma configuração incorreta do padrão de nome de arquivo nas propriedades do Log4j2. A configuração incorreta faz com que os componentes gerem somente um arquivo de log por dia. Quando a estratégia de rotação ocorre, ela substitui o arquivo existente em vez de gerar um novo arquivo de log, conforme esperado. Como solução alternativa, use uma ação de bootstrap para gerar arquivos de log a cada hora e acrescentar um número inteiro de incremento automático no nome do arquivo para lidar com a rotação.

  Para as versões 6.6.0 a 6.9.0 do Amazon EMR, use a seguinte ação de bootstrap ao iniciar um cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

  Para o Amazon EMR 5.36.0, use a ação de bootstrap a seguir ao iniciar um cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

Para obter mais informações sobre o cronograma da versão, consulte o log de alterações.

Versão 6.7.0

As notas da versão a seguir incluem informações para a versão 6.7.0 do Amazon EMR. As alterações são referentes à versão 6.6.0.

Data da versão inicial: 15 de julho de 2022

Novos atributos

- O Amazon EMR agora é compatível com Apache Spark 3.2.1, Apache Hive 3.1.3, HUDI 0.11, PrestoDB 0.272 e Trino 0.378.
- Suporta controles de acesso baseados em IAM Role e Lake Formation com etapas do EMR (Spark, Hive) para Amazon EMR em clusters. EC2
- Ele é compatível com instruções de definição de dados do Apache Spark em clusters habilitados para Apache Ranger. Isso agora inclui suporte para aplicações do Trino lendo e gravando metadados do Apache Hive em clusters habilitados para Apache Ranger. Para obter mais informações, consulte Habilitar governança federada usando Trino e Apache Ranger no Amazon EMR.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			, China (Pequim), Ásia-Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda), Oriente Médio (Emirados Árabes Unidos), Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Hyderaba d), Leste dos EUA (Norte da Virgínia), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Europa (Londres) , Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA), Oeste dos EUA (Oregon) AWS GovCloud , Oeste dos EUA (Norte da Califórnia) AWS GovCloud , Ásia-Pacífico (Cingapura), Ásia-Pací fico (Sydney)), Israel (Tel Aviv), Canadá (Central) , Europa (Espanha), China (Ningxia), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Espanha), Europa (Estocolmo), Europa (Zurique), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Hyderabad), Ásia-Pacífico (Jacarta) Ásia-Pacífico (Mumbai), Ásia-Pacífico (Osaka), Ásia-Pacífico (Seul), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (EAU), Canadá (Central) , (EUA- AWS GovCloud Leste), AWS GovCloud (Oeste dos EUA), China

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			(Pequim), China (Ningxia), Israel (Tel Aviv)
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Centro Leste (Emirados Árabes Unidos), Israel (Tel Aviv), Europa (Zurique)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 001.0	4.14.352	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 816.0	4.14.350	21 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 809.0	4.14.349	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 719.0	4.14.348	25 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia AWS GovCloud), Ásia Pacífico (Hyderaba d), Oriente Médio (EAU),

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Europa (Espanha), Europa (Zurique)
2.0.2024 223.0	4.14.336	8 de março de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 131.0	4.14.336	14 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 124.0	4.14.336	7 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 109.0	4.14.334	24 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 218.0	4.14.330	2 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 206.0	4.14.330	22 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 116.0	4.14.328	11 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 101.0	4.14.327	16 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 020.1	4.14.326	7 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 012.1	4.14.326	26 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 926.0	4.14.322	19 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 906.0	4.14.322	4 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 822.0	4.14.322	30 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 808.0	4.14.320	24 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 727.0	4.14.320	14 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 719.0	4.14.320	2 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta) , África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 628.0	4.14.318	12 de julho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 612.0	4.14.314	23 de junho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 504.1	4.14.313	16 de maio de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 418.0	4.14.311	3 de maio de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 404.1	4.14.311	18 de abril de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)
2.0.2023 404.0	4.14.311	10 de abril de 2023	Leste dos EUA (Norte da Virgínia), Europa (Paris)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 320.0	4.14.309	30 de março de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 307.0	4.14.305	15 de março de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 207.0	4.14.304	22 de fevereiro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 119.1	4.14.301	3 de fevereiro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 210.1	4.14.301	22 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 103.3	4.14.296	5 de dezembro de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 004.0	4.14.294	2 de novembro de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 912.1	4.14.291	7 de outubro de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)
2.0.2022 719.0	4.14.287	10 de agosto de 2022	`us-west-1` , `eu-west-3` , `eu-north-1` , `ap-south-1` , `me-south-1`

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 606.1	4.14.281	15 de julho de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

Problemas conhecidos

- Quando a versão 6.5.0, 6.6.0 ou 6.7.0 do Amazon EMR lê as tabelas do Apache Phoenix por meio do shell do Apache Spark, ocorre um `NoSuchMethodError` porque o Amazon EMR usa um `Hbase.compat.version` incorreto. A versão 6.8.0 do Amazon EMR corrige esse problema.

- Ao usar o conector DynamoDB com o Spark nas versões 6.6.0, 6.7.0 e 6.8.0 do Amazon EMR, todas as leituras da tabela retornam um resultado vazio, mesmo que a divisão de entrada faça referência a dados que não estão vazios. Isso ocorre porque o Spark 3.2.0 define `spark.hadoopRDD.ignoreEmptySplits` como `true` por padrão. Como solução alternativa, defina explicitamente `spark.hadoopRDD.ignoreEmptySplits` como `false`. A versão 6.9.0 do Amazon EMR corrige esse problema.
- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
  - O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

  Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.
- Com as versões 5.36.0 e 6.6.0 a 6.9.0 do Amazon EMR, os componentes do serviço `SecretAgent` e `RecordServer` podem sofrer perda de dados de log devido a uma configuração incorreta do padrão de nome de arquivo nas propriedades do Log4j2. A configuração incorreta faz com que os componentes gerem somente um arquivo de log por dia. Quando a estratégia de rotação ocorre, ela substitui o arquivo existente em vez de gerar um novo arquivo de log, conforme esperado. Como solução alternativa, use uma ação de bootstrap para gerar arquivos de log a cada hora e acrescentar um número inteiro de incremento automático no nome do arquivo para lidar com a rotação.

  Para as versões 6.6.0 a 6.9.0 do Amazon EMR, use a seguinte ação de bootstrap ao iniciar um cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

Para o Amazon EMR 5.36.0, use a ação de bootstrap a seguir ao iniciar um cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

- A API `GetClusterSessionCredentials` não é compatível com clusters executados no Amazon EMR 6.7 ou versões inferiores.

Versão 6.6.0

As notas da versão a seguir incluem informações para a versão 6.6.0 do Amazon EMR. As alterações são referentes à versão 6.5.0.

Data da versão inicial: 9 de maio de 2022

Data da documentação atualizada: 15 de junho de 2022

Novos atributos

- O Amazon EMR 6.6 agora é compatível com Apache Spark 3.2, Apache Spark RAPIDS 22.02, CUDA 11, Apache Hudi 0.10.1, Apache Iceberg 0.13, Trino 0.367 e PrestoDB 0.267.
- Quando você inicia um cluster com a versão de patch mais recente do Amazon EMR 5.36 ou superior, 6.6 ou superior ou 7.0 ou superior, o Amazon EMR usa a versão mais recente do Amazon Linux 2023 ou do Amazon Linux 2 para a AMI padrão do Amazon EMR. Para obter mais informações, consulte Como usar a AMI padrão do Amazon Linux para Amazon EMR.

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junho de 2025	Europa (Estocolmo), Oriente Médio (Bahrein) , China (Pequim), Ásia-

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
			Pacífico (Mumbai), Europa (Paris), Ásia-Pacífico (Jacarta), Leste dos EUA (Ohio), África (Cidade do Cabo), Europa (Irlanda) , Europa (Frankfurt), América do Sul (São Paulo), Ásia-Pacífico (Hong Kong), Leste dos EUA (Norte da Virgínia) , Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Europa (Londres), Europa (Milão), Ásia-Pacífico (Tóquio), (Leste dos EUA), (Oeste dos EUA) AWS GovCloud , Oeste dos EUA (Oregon) AWS GovCloud , Oeste dos EUA (Norte da Califórnia), Ásia-Pacífico (Cingapur a), Ásia-Pacífico (Sydney), Canadá (Central), China (Ningxia)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 512.0	4.14.355-277.643.amzn2	04 de junho de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Milão), Europa (Paris), Europa (Estocolmo), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Jacarta), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Osaka), Ásia-Pací fico (Seul)), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Tóquio), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (Leste dos EUA), (Oeste dos AWS GovCloud EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de maio de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfurt), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 321.0	4.14.355	09 de abril de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 305.0	4.14.355	18 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 220.0	4.14.355	08 de março de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 201.0	4.14.355	28 de fevereiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 123.4	4.14.355	27 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2025 116.0	4.14.355	23 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 217.0	4.14.355	8 de janeiro de 2025	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 001.0	4.14.352	4 de outubro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 816.0	4.14.350	21 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 809.0	4.14.349	20 de agosto de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 719.0	4.14.348	25 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 709.1	4.14.348	23 de julho de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 223.0	4.14.336	8 de março de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 131.0	4.14.336	14 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 124.0	4.14.336	7 de fevereiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2024 109.0	4.14.334	24 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 218.0	4.14.330	2 de janeiro de 2024	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 206.0	4.14.330	22 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 116.0	4.14.328	11 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 101.0	4.14.327	16 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 020.1	4.14.326	7 de novembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 012.1	4.14.326	26 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 926.0	4.14.322	19 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Ásia-Pacífico (Osaka), Ásia-Pacífico (Cingapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central), (AWS GovCloud Oeste dos EUA), (Leste dos EUA), China (Pequim), China (Ningxia) AWS GovCloud

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 906.0	4.14.322	4 de outubro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 822.0	4.14.322	30 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 808.0	4.14.320	24 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 727.0	4.14.320	14 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Frankfur t), Europa (Irlanda), Europa (Londres), Europa (Paris), Ásia-Pacífico (Hong Kong), Ásia-Pací fico (Mumbai), Ásia-Pací fico (Tóquio), Ásia-Pací fico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 719.0	4.14.320	2 de agosto de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Europa (Estocolmo), Europa (Milão), Europa (Espanha), Europa (Frankfurt), Europa (Zurique), Europa (Irlanda) , Europa (Londres), Europa (Paris), Ásia-Pací fico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Hyderabad), Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), Ásia-Pacífico (Jacarta) , África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein), Oriente Médio (Emirados Árabes Unidos), Canadá (Central)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 628.0	4.14.318	12 de julho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 612.0	4.14.314	23 de junho de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 504.1	4.14.313	16 de maio de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 418.0	4.14.311	3 de maio de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 404.1	4.14.311	18 de abril de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)
2.0.2023 404.0	4.14.311	10 de abril de 2023	Leste dos EUA (Norte da Virgínia), Europa (Paris)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 320.0	4.14.309	30 de março de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 307.0	4.14.305	15 de março de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 207.0	4.14.304	22 de fevereiro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2023 119.1	4.14.301	3 de fevereiro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 210.1	4.14.301	22 de dezembro de 2023	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 103.3	4.14.296	5 de dezembro de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 004.0	4.14.294	2 de novembro de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 912.1	4.14.291	7 de outubro de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)
2.0.2022 805.0	4.14.287	30 de agosto de 2022	`us-west-1`

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 719.0	4.14.287	10 de agosto de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 426.0	4.14.281	10 de junho de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

OsRelea Label (Versão Amazon Linux)	Versão do kernel do Amazon Linux	Data disponível	Regiões do compatíveis
2.0.2022 406.1	4.14.275	2 de maio de 2022	Leste dos EUA (Norte da Virgínia), Leste dos EUA (Ohio), Oeste dos EUA (Norte da Califórnia), Oeste dos EUA (Oregon), Canadá (Central), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (Paris), Europa (Frankfurt), Europa (Milão), Ásia-Pacífico (Hong Kong), Ásia-Pacífico (Mumbai), Ásia-Pacífico (Jacarta) , Ásia-Pacífico (Tóquio), Ásia-Pacífico (Seul), Asia Pacific (Osaka), Ásia-Pacífico (Singapura), Ásia-Pacífico (Sydney), África (Cidade do Cabo), América do Sul (São Paulo), Oriente Médio (Bahrein)

- Com as versões 6.6 e posteriores do Amazon EMR,as aplicações que usam o Log4j 1.x e o Log4j 2.x são atualizadas para usar o Log4j 1.2.17 (ou superior) e o Log4j 2.17.1 (ou superior), respectivamente, e não exigem o uso de ações de bootstrap para mitigar os problemas de CVE.
- [Ajuste de escala gerenciado] Otimização do ajuste de escala gerenciado de dados embaralhados do Spark: para as versões 5.34.0 e posteriores do Amazon EMR e as versões 6.4.0 e posteriores do EMR, o ajuste de escala gerenciado agora reconhece dados embaralhados do Spark (dados que o Spark redistribui entre partições para executar operações específicas). Para obter mais

informações sobre operações de shuffle, consulte Usar ajuste de escala gerenciado do EMR no Amazon EMR no Guia de gerenciamento do Amazon EMR e no Guia de programação do Spark.

- Desde as versões 5.32.0 e 6.5.0 do Amazon EMR, o dimensionamento do executor dinâmico para o Apache Spark está habilitado por padrão. Para ativar ou desativar esse atributo, você pode usar o parâmetro de configuração `spark.yarn.heterogeneousExecutors.enabled`.

Alterações, melhorias e problemas resolvidos

- O Amazon EMR reduz o tempo de inicialização do cluster em até 80 segundos, em média, para clusters que usam a opção de AMI padrão do EMR e só instalam aplicações comuns, como Apache Hadoop, Apache Spark e Apache Hive.

Problemas conhecidos

- Quando a versão 6.5.0, 6.6.0 ou 6.7.0 do Amazon EMR lê as tabelas do Apache Phoenix por meio do shell do Apache Spark, ocorre um `NoSuchMethodError` porque o Amazon EMR usa um `Hbase.compat.version` incorreto. A versão 6.8.0 do Amazon EMR corrige esse problema.
- Ao usar o conector DynamoDB com o Spark nas versões 6.6.0, 6.7.0 e 6.8.0 do Amazon EMR, todas as leituras da tabela retornam um resultado vazio, mesmo que a divisão de entrada faça referência a dados que não estão vazios. Isso ocorre porque o Spark 3.2.0 define `spark.hadoopRDD.ignoreEmptySplits` como `true` por padrão. Como solução alternativa, defina explicitamente `spark.hadoopRDD.ignoreEmptySplits` como `false`. A versão 6.9.0 do Amazon EMR corrige esse problema.
- Em clusters de execução prolongada do Trino, o Amazon EMR 6.6.0 habilita os parâmetros de registro em log da coleta de resíduos no jvm.config do Trino para obter melhores insights dos registros em log da coleta de resíduos. Essa alteração anexa muitos registros da coleta de lixo ao arquivo launcher.log (/var/log/trino/launcher.log). Se você estiver executando clusters do Trino no Amazon EMR 6.6.0, poderá encontrar nós que perderão espaço em disco depois que o cluster estiver em execução por alguns dias devido aos logs anexados.

  A solução alternativa para esse problema é executar o script abaixo como uma ação de bootstrap para desativar os parâmetros de registro em log da coleta de resíduos no jvm.config ao criar ou clonar o cluster para o Amazon EMR 6.6.0.

```
#!/bin/bash
  set -ex
  PRESTO_PUPPET_DIR='/var/aws/emr/bigtop-deploy/puppet/modules/trino'
```

```
  sudo bash -c "sed -i '/-Xlog/d' ${PRESTO_PUPPET_DIR}/templates/jvm.config"
```

- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
  - O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

  Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.
- Com as versões 5.36.0 e 6.6.0 a 6.9.0 do Amazon EMR, os componentes do serviço `SecretAgent` e `RecordServer` podem sofrer perda de dados de log devido a uma configuração incorreta do padrão de nome de arquivo nas propriedades do Log4j2. A configuração incorreta faz com que os componentes gerem somente um arquivo de log por dia. Quando a estratégia de rotação ocorre, ela substitui o arquivo existente em vez de gerar um novo arquivo de log, conforme esperado. Como solução alternativa, use uma ação de bootstrap para gerar arquivos de log a cada hora e acrescentar um número inteiro de incremento automático no nome do arquivo para lidar com a rotação.

  Para as versões 6.6.0 a 6.9.0 do Amazon EMR, use a seguinte ação de bootstrap ao iniciar um cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

  Para o Amazon EMR 5.36.0, use a ação de bootstrap a seguir ao iniciar um cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

Versão 5.35.0

Esta é a nota de lançamento da versão 5.35.0 do Amazon EMR.

As notas da versão a seguir incluem informações para a versão 5.35.0 do Amazon EMR. As alterações são referentes à versão 5.34.0.

Data da versão inicial: 30 de março de 2022

Novos atributos

- As aplicações da versão 5.35 do Amazon EMR que usam o Log4j 1.x e o Log4j 2.x são atualizadas para usar o Log4j 1.2.17 (ou superior) e o Log4j 2.17.1 (ou superior), respectivamente, e não exigem o uso de ações de bootstrap para mitigar os problemas de CVE em versões anteriores. Consulte Abordagem para mitigar o CVE-2021-44228.

Alterações, melhorias e problemas resolvidos

Alterações no Flink

Alterar tipo	Descrição
Atualizações	• Atualizar a versão do flink para 1.14.2. • log4j atualizado para 2.17.1.

Alterações no Hadoop

Alterar tipo	Descrição
Backports de código aberto do Hadoop desde o EMR 5.34.0	• YARN-10438: Manipule o ContainerID nulo no cliente # () RMService getContainerReport

Alterar tipo	Descrição
	• YARN-7266: threads do manipulador de eventos do servidor de linha do tempo bloqueados • YARN-10438: O ATS 1.5 falha ao iniciar se os RollingLevelDb arquivos estiverem corrompidos ou ausentes • HADOOP-13500: sincronizar a iteração do objeto de propriedades de configuração • YARN-10651: CapacityScheduler travou com a entrada NPE. AbstractYarnScheduler updateNodeResource() • HDFS-12221: Substitua o xerces em XmlEditsVisitor • HDFS-16410: Análise insegura de Xml em OfflineEditsXmlLoader
Alterações e correções no Hadoop	• O Tomcat usado em KMS e HttpFS foi atualizado para 8.5.75 • Na FileSystemOptimizedCommitter V2, o marcador de sucesso foi escrito no caminho de saída do commitJob definido durante a criação do committer. Como os caminhos de saída do commitJob e do nível da tarefa podem ser diferentes, o caminho foi corrigido para usar o caminho definido nos arquivos de manifesto. Para trabalhos do Hive, isso faz com que o marcador de sucesso seja gravado corretamente ao serem executada s operações, como partição dinâmica ou UNION ALL.

Alterações no Hive

Alterar tipo	Descrição
O Hive foi atualizado para a versão 2.3.9 de código aberto, incluindo estas correções do JIRA	• HIVE-17155: findConfFile () em HiveConf .java tem alguns problemas com o caminho de configuração • HIVE-24797: desabilitar a validação de valores padrão ao quando analisar esquemas Avro • HIVE-21563: Melhore o desempenho do getEmptyTable Table# desativando o Once registerAllFunctions • HIVE-18147: Os testes podem falhar com o java.net. BindException: Endereço já em uso • HIVE-24608: voltar para get_table no cliente HMS para Hive 2.3.x • HIVE-21200: Vetorização - coluna de data lançando java.lang. UnsupportedOperati onException para parquet • HIVE-19228: remover o uso do commons-h ttpclient 3.x
Backports de código aberto do Hive desde o EMR 5.34.0	• HIVE-19990: falha na consulta com intervalo literal na condição de junção • HIVE-25824: atualizar o branch-2.3 para log4j 2.17.0 • TEZ-4062: o agendamento de tentativas especulativas deverá ser anulado quando a tarefa for concluída • TEZ-4108: NullPointerException durante a execução especulativa, condição de corrida • TEZ-3918: a configuração de tez.task. log.level não funciona
Atualizações e correções do Hive	• Atualizar a versão Log4j para 2.17.1

Alterar tipo	Descrição
	• Atualizar a versão do ORC para 1.4.3 • Corrigido o impasse devido à discussão de penalidades em ShuffleScheduler
Novos recursos	• Adicionado atributo para imprimir consultas do Hive em logs AM. Ela fica desabilitada por padrão. Sinalizador/Conf.: `tez.am.emr.print.hive.query.in.log` . Status (padrão): FALSO.

Alterações no Oozie

Alterar tipo	Descrição
Backports de código aberto do Oozie desde o EMR 5.34.0	• OOZIE-3652: O lançador Oozie deve tentar novamente a listagem de diretórios quando ocorrer NoSuchFileException

Alterações no Pig

Alterar tipo	Descrição
Atualizações	• log4j atualizado para 1.2.17.

Problemas conhecidos

- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.

- O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere / (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.

Versão 5.34.0

As notas da versão a seguir incluem informações para a versão 5.34.0 do Amazon EMR. As alterações são referentes à versão 5.33.1.

Data da versão inicial: 20 de janeiro de 2022

Data da versão atualizada: 21 de março de 2022

Novos atributos

- [Ajuste de escala gerenciado] Otimização do ajuste de escala gerenciado de dados embaralhados do Spark: para as versões 5.34.0 e posteriores do Amazon EMR e as versões 6.4.0 e posteriores do EMR, o ajuste de escala gerenciado agora reconhece dados embaralhados do Spark (dados que o Spark redistribui entre partições para executar operações específicas). Para obter mais informações sobre operações de shuffle, consulte Usar ajuste de escala gerenciado do EMR no Amazon EMR no Guia de gerenciamento do Amazon EMR e no Guia de programação do Spark.
- [Hudi] Melhorias para simplificar a configuração do Hudi. Desabilitado o controle de simultaneidade otimista por padrão.

Alterações, melhorias e problemas resolvidos

- Esta é uma versão para corrigir problemas com o Amazon EMR Scaling quando ele não consegue up/scale reduzir a escala de um cluster com sucesso ou causa falhas no aplicativo.
- Anteriormente, a reinicialização manual do gerenciador de recursos em um cluster multimestre fazia com que os daemons do Amazon EMR no cluster, como o Zookeeper, recarregassem todos os nós anteriormente desativados ou perdidos no arquivo znode do Zookeeper. Isso fez com que

os limites padrão fossem excedidos em determinadas situações. O Amazon EMR agora remove os registros de nós desativados ou perdidos há mais de uma hora do arquivo do Zookeeper e os limites internos foram aumentados.

- Corrigido um problema em que as solicitações de escalabilidade falhavam em um cluster grande e altamente utilizado quando os daemons do Amazon EMR no cluster estavam executando atividades de verificação de integridade, como a coleta do estado do nó do YARN e o estado do nó do HDFS. Isso estava acontecendo porque os daemons no cluster não conseguiam comunicar os dados do status de integridade de um nó aos componentes internos do Amazon EMR.
- Aprimorados os daemons do EMR no cluster para rastrear corretamente os estados dos nós quando são reutilizados endereços IP para melhorar a confiabilidade durante operações de escalabilidade.
- SPARK-29683. Corrigido um problema em que ocorriam falhas de trabalho durante a redução da escala verticalmente do cluster, pois o Spark presumia que todos os nós disponíveis estavam na lista de negação.
- YARN-9011. Corrigido um problema em que ocorriam falhas de trabalho devido a uma condição de corrida na desativação do YARN quando o cluster tentava aumentar ou reduzir a escala verticalmente.
- Corrigido problema com falhas de etapas ou tarefas durante a escalabilidade do cluster ao ser garantido que os estados dos nós fossem sempre consistentes entre os daemons do Amazon EMR no cluster e o YARN/HDFS.
- Corrigido um problema em que operações de cluster, como redução de escala verticalmente e envio de etapas, falhavam para clusters do Amazon EMR habilitados com a autenticação Kerberos. Isso ocorreu porque o daemon no cluster do Amazon EMR não renovou o tíquete Kerberos, que é necessário para se comunicar com segurança com a execução no nó primário. HDFS/YARN
- Zeppelin atualizado para a versão 0.10.0.
- Livy Fix: atualizado para 0.7.1
- Melhoria da performance do Spark: executores heterogêneos são desabilitados quando determinados valores de configuração do Spark são substituídos no EMR 5.34.0.
- WebHDFS e o servidor HttpFS estão desabilitados por padrão. Você pode reabilitar o WebHDFS usando a configuração do Hadoop, `dfs.webhdfs.enabled`. O servidor HttpFS pode ser iniciado usando `sudo systemctl start hadoop-httpfs`.

Problemas conhecidos

- O atributo Cadernos do Amazon EMR usado com a personificação de usuários do Livy não funciona porque o HttpFS está desabilitado por padrão. Nesse caso, o caderno do EMR não pode se conectar ao cluster que tem a personificação do Livy habilitada. A solução alternativa é iniciar o servidor HttpFS antes de conectar o caderno do EMR ao cluster usando `sudo systemctl start hadoop-httpfs`.
- As consultas no Hue não funcionam no Amazon EMR 6.4.0 porque o servidor HttpFS do Apache Hadoop está desabilitado por padrão. Para usar o Hue no Amazon EMR 6.4.0, inicie manualmente o servidor HttpFS no nó primário do Amazon EMR usando `sudo systemctl start hadoop-httpfs`, ou use uma etapa do Amazon EMR.
- O atributo Cadernos do Amazon EMR usado com a personificação de usuários do Livy não funciona porque o HttpFS está desabilitado por padrão. Nesse caso, o caderno do EMR não pode se conectar ao cluster que tem a personificação do Livy habilitada. A solução alternativa é iniciar o servidor HttpFS antes de conectar o caderno do EMR ao cluster usando `sudo systemctl start hadoop-httpfs`.
- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
  - O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

  Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.

Versão 6.5.0

As notas da versão a seguir incluem informações para a versão 6.5.0 do Amazon EMR. As alterações são referentes à versão 6.4.0.

Data da versão inicial: 20 de janeiro de 2022

Data da versão atualizada: 21 de março de 2022

Novos atributos

- [Ajuste de escala gerenciado] Otimização do ajuste de escala gerenciado de dados embaralhados do Spark: para as versões 5.34.0 e posteriores do Amazon EMR e as versões 6.4.0 e posteriores do EMR, o ajuste de escala gerenciado agora reconhece dados embaralhados do Spark (dados que o Spark redistribui entre partições para executar operações específicas). Para obter mais informações sobre operações de shuffle, consulte Usar ajuste de escala gerenciado do EMR no Amazon EMR no Guia de gerenciamento do Amazon EMR e no Guia de programação do Spark.
- Desde as versões 5.32.0 e 6.5.0 do Amazon EMR, o dimensionamento do executor dinâmico para o Apache Spark está habilitado por padrão. Para ativar ou desativar esse atributo, você pode usar o parâmetro de configuração `spark.yarn.heterogeneousExecutors.enabled`.
- Suporte para o formato de tabela aberta Apache Iceberg para conjuntos de dados analíticos imensos.
- Support para ranger-trino-plugin 2.0.1-amzn-1
- Suporte para toree 0.5.0

Alterações, melhorias e problemas resolvidos

- A versão de lançamento do Amazon EMR 6.5 agora é compatível com o Apache Iceberg 0.12.0 e fornece melhorias no runtime com o Ambiente de Tempo de Execução do Amazon EMR para Apache Spark, o Ambiente de Tempo de Execução do Amazon EMR para Presto e o Runtime do Amazon EMR para Apache Hive.
- O Apache Iceberg é um formato de tabela aberta para grandes conjuntos de dados no Amazon S3 que fornece performance rápida de consultas em tabelas grandes, confirmações atômicas, gravações simultâneas e evolução de tabelas compatível com SQL. Com o EMR 6.5, você pode usar o Apache Spark 3.1.2 com o formato de tabela Iceberg.

- O Apache Hudi 0.9 adiciona suporte a DDL e DML do Spark SQL. Isso permite a criação e a atualização de tabelas do Hudi usando apenas instruções de SQL. O Apache Hudi 0.9 também inclui melhorias na performance do lado da consulta e do lado do gravador.
- O Runtime do Amazon EMR para Apache Hive melhora a performance do Apache Hive no Amazon S3 ao remover operações de renomeação durante operações de preparação e melhora a performance dos comandos de verificação do metastore (MSCK) usados para reparar tabelas.

Problemas conhecidos

- Quando a versão 6.5.0, 6.6.0 ou 6.7.0 do Amazon EMR lê as tabelas do Apache Phoenix por meio do shell do Apache Spark, ocorre um `NoSuchMethodError` porque o Amazon EMR usa um `Hbase.compat.version` incorreto. A versão 6.8.0 do Amazon EMR corrige esse problema.
- Os clusters do pacote do Hbase em alta disponibilidade (HA) apresentam falha no provisionamento com o tamanho de volume e o tipo de instância padrão. A solução alternativa para esse problema é aumentar o tamanho do volume raiz.
- Para usar as ações do Spark com o Apache Oozie, você deve adicionar a seguinte configuração ao seu arquivo `workflow.xml` do Oozie. Caso contrário, várias bibliotecas críticas, como Hadoop e EMRFS, estarão ausentes do classpath dos executores do Spark que o Oozie inicia.

```
<spark-opts>--conf spark.yarn.populateHadoopClasspath=true</spark-opts>
```

- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
  - O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.

Versão 6.4.0

As notas da versão a seguir incluem informações para a versão 6.4.0 do Amazon EMR. As alterações são referentes à versão 6.3.0.

Data da versão inicial: 20 de setembro de 2021

Data da versão atualizada: 21 de março de 2022

Aplicações compatíveis

- AWS SDK para Java versão 1.12.31
- CloudWatch Sink versão 2.2.0
- Conector do DynamoDB versão 4.16.0
- EMRFS versão 2.47.0
- Amazon EMR Goodies versão 3.2.0
- Conector Kinesis versão 3.5.0 do Amazon EMR
- Servidor de registros do Amazon EMR versão 2.1.0
- Scripts versão 2.5.0 do Amazon EMR
- Flink versão 1.13.1
- Ganglia versão 3.7.2
- AWS Glue Hive Metastore Client versão 3.3.0
- Hadoop versão 3.2.1-amzn-4
- HBase versão 2.4.4-amzn-0
- HBase-operator-tools 1.1.0
- HCatalog versão 3.1.2-amzn-5
- Hive versão 3.1.2-amzn-5
- Hudi versão 0.8.0-amzn-0
- Hue versão 4.9.0
- Java JDK versão Corretto-8.302.08.1 (compilação 1.8.0_302-b08)

- JupyterHub versão 1.4.1
- Livy versão 0.7.1-incubating
- MXNet versão 1.8.0
- Oozie versão 5.2.1
- Phoenix versão 5.1.2
- Pig versão 0.17.0
- Presto versão 0.254.1-amzn-0
- Trino versão 359
- KMS do Apache Ranger (criptografia transparente multi-mestre) versão 2.0.0
- ranger-plugins 2.0.1-amzn-0
- ranger-s3-plugin 1.2.0
- SageMaker SDK do Spark versão 1.4.1
- Scala versão 2.12.10 (VM de servidor OpenJDK de 64 bits, Java 1.8.0_282)
- Spark versão 3.1.2-amzn-0
- spark-rapids 0.4.1
- Sqoop versão 1.4.7
- TensorFlow versão 2.4.1
- Tez versão 0.9.2
- Zeppelin versão 0.9.0
- Zookeeper versão 3.5.7
- Conectores e drivers: DynamoDB Connector 4.16.0

Novos recursos

- [Ajuste de escala gerenciado] Otimização do ajuste de escala gerenciado de dados embaralhados do Spark: para as versões 5.34.0 e posteriores do Amazon EMR e as versões 6.4.0 e posteriores do EMR, o ajuste de escala gerenciado agora reconhece dados embaralhados do Spark (dados que o Spark redistribui entre partições para executar operações específicas). Para obter mais informações sobre operações de shuffle, consulte Usar ajuste de escala gerenciado do EMR no Amazon EMR no Guia de gerenciamento do Amazon EMR e no Guia de programação do Spark.
- Em clusters do Amazon EMR habilitados para Apache Ranger, você pode usar o Apache Spark SQL para inserir dados ou atualizar as tabelas de metastore do Apache Hive usando `INSERT`

`INTO`, `INSERT OVERWRITE` e `ALTER TABLE`. Ao ser usado ALTER TABLE com Spark SQL, o local da partição deve ser o diretório filho do local de uma tabela. Atualmente,o Amazon EMR não permite a inserção de dados em uma partição cuja localização seja diferente da localização da tabela.

- O PrestoSQL foi renomeado como Trino.
- Hive: a execução de consultas SELECT simples com a cláusula LIMIT é acelerada com a interrupção da execução da consulta assim que o número de registros mencionados na cláusula LIMIT é obtido. As consultas SELECT simples são consultas que não têm a cláusula GROUP BY/ORDER by ou consultas que não têm um estágio redutor. Por exemplo, `.SELECT * from <TABLE> WHERE <Condition> LIMIT <Number>`

Controle de simultaneidade do Hudi

- O Hudi, agora, é compatível com o Optimistic Concurrency Control (OCC - Controle de simultaneidade otimista), que pode ser aproveitado com operações de gravação como UPSERT e INSERT para permitir alterações de vários gravadores na mesma tabela do Hudi. Esse é o OCC em nível de arquivo. Portanto, quaisquer duas confirmações (ou gravadores) podem gravar na mesma tabela, se suas alterações não são conflitantes. Para obter mais informações, consulte o Controle de simultaneidade do Hudi.
- Os clusters do Amazon EMR têm o Zookeeper instalado, que pode ser usado como provedor de bloqueio para o OCC. Para facilitar o uso desse atributo, os clusters do Amazon EMR têm as seguintes propriedades pré-configuradas:

```
hoodie.write.lock.provider=org.apache.hudi.client.transaction.lock.ZookeeperBasedLockProvider
hoodie.write.lock.zookeeper.url=<EMR Zookeeper URL>
hoodie.write.lock.zookeeper.port=<EMR Zookeeper Port>
hoodie.write.lock.zookeeper.base_path=/hudi
```

Para habilitar o OCC, você precisa configurar as seguintes propriedades com as opções de trabalho do Hudi ou em nível de cluster usando a API de configurações do Amazon EMR:

```
hoodie.write.concurrency.mode=optimistic_concurrency_control
hoodie.cleaner.policy.failed.writes=LAZY (Performs cleaning of failed writes lazily
 instead of inline with every write)
hoodie.write.lock.zookeeper.lock_key=<Key to uniquely identify the Hudi table> (Table
 Name is a good option)
```

Monitoramento Hudi: CloudWatch integração com a Amazon para reportar Hudi Metrics

- O Amazon EMR oferece suporte à publicação de Hudi Metrics na Amazon. CloudWatch Isso é habilitado com a definição das seguintes configurações necessárias:

```
hoodie.metrics.on=true
hoodie.metrics.reporter.type=CLOUDWATCH
```

- A seguir, são mostradas as configurações opcionais do Hudi que você pode alterar:

Configuração	Descrição	Valor
hoodie.metrics.cloudwatch.report.period.seconds	Frequência (em segundos) na qual reportar métricas para a Amazon CloudWatch	O valor padrão é 60s, o que é adequado para a resolução padrão de um minuto oferecida pela Amazon CloudWatch
hoodie.metrics.cloudwatch.metric.prefix	Prefixo a ser adicionado ao nome de cada métrica	O valor padrão está vazio (sem prefixo)
hoodie.metrics.cloudwatch.namespace	CloudWatch Namespace da Amazon sob o qual as métricas são publicadas	O valor padrão é Hudi
hoodie.metrics.cloudwatch.maxDatumsPerSolicitação	Número máximo de dados a serem incluídos em uma solicitação à Amazon CloudWatch	O valor padrão é 20, que é o mesmo que o CloudWatch padrão da Amazon

Suporte e melhorias das configurações do Hudi no Amazon EMR

- Agora, os clientes podem aproveitar a API de configurações e o atributo de reconfiguração do EMR para definir as configurações do Hudi em nível de cluster. Um novo suporte de configuração baseado em arquivos foi introduzido via/etc/hudi/conf/hudi-defaults.conf nos moldes de outros aplicativos como Spark, Hive etc. O EMR configura alguns padrões para melhorar a experiência do usuário:

— `hoodie.datasource.hive_sync.jdbcurl` está configurado para a URL do servidor do Hive do cluster e não precisa mais ser especificado. Isso é particularmente útil na execução de um trabalho no modo de cluster do Spark, em que anteriormente era necessário especificar o IP principal do Amazon EMR.

— configurações HBase específicas, que são úteis para usar o HBase índice com o Hudi.

— Configuração específica do provedor de bloqueio do Zookeeper, conforme discutido em Controle de simultaneidade, o que facilita o uso do Controle de simultaneidade otimista (OCC).

- Alterações adicionais foram introduzidas para reduzir o número de configurações que você precisa passar e inferir automaticamente sempre que possível:

  — A palavra-chave `partitionBy` pode ser usada para especificar a coluna de partição.

  — Ao habilitar o Hive Sync, não é mais obrigatório passar `HIVE_TABLE_OPT_KEY, HIVE_PARTITION_FIELDS_OPT_KEY, HIVE_PARTITION_EXTRACTOR_CLASS_OPT_KEY`. Esses valores podem ser deduzidos com base no nome da tabela Hudi e no campo de partição.

  — não é obrigatório passar `KEYGENERATOR_CLASS_OPT_KEY`, que pode ser inferido com base em casos mais simples de `SimpleKeyGenerator` e `ComplexKeyGenerator`.

Advertências do Hudi

- O Hudi não permite execução vetorizada no Hive de tabelas Merge on Read (MoR - Mesclar na leitura) e Bootstrap. Por exemplo, `count(*)` apresenta falha com a tabela do Hudi em tempo real quando `hive.vectorized.execution.enabled` está definido como verdadeiro. Como solução alternativa, você pode desabilitar a leitura vetorizada configurando `hive.vectorized.execution.enabled` como `false`.
- O suporte a vários gravadores não é compatível com o atributo de bootstrap do Hudi.
- O streamer e o SQL do Flink são atributos experimentais nesta versão. Esses atributos não são recomendados para uso em implantações de produção.

Alterações, melhorias e problemas resolvidos

Esta é uma versão para corrigir problemas com o Amazon EMR Scaling quando ele não consegue up/scale reduzir a escala de um cluster com sucesso ou causa falhas no aplicativo.

- Anteriormente, a reinicialização manual do gerenciador de recursos em um cluster multimestre fazia com que os daemons do Amazon EMR no cluster, como o Zookeeper, recarregassem todos os nós anteriormente desativados ou perdidos no arquivo znode do Zookeeper. Isso fez com que os limites padrão fossem excedidos em determinadas situações. O Amazon EMR agora remove os registros de nós desativados ou perdidos há mais de uma hora do arquivo do Zookeeper e os limites internos foram aumentados.
- Corrigido um problema em que as solicitações de escalabilidade falhavam em um cluster grande e altamente utilizado quando os daemons do Amazon EMR no cluster estavam executando atividades de verificação de integridade, como a coleta do estado do nó do YARN e o estado do nó do HDFS. Isso estava acontecendo porque os daemons no cluster não conseguiam comunicar os dados do status de integridade de um nó aos componentes internos do Amazon EMR.
- Aprimorados os daemons do EMR no cluster para rastrear corretamente os estados dos nós quando são reutilizados endereços IP para melhorar a confiabilidade durante operações de escalabilidade.
- SPARK-29683. Corrigido um problema em que ocorriam falhas de trabalho durante a redução da escala verticalmente do cluster, pois o Spark presumia que todos os nós disponíveis estavam na lista de negação.
- YARN-9011. Corrigido um problema em que ocorriam falhas de trabalho devido a uma condição de corrida na desativação do YARN quando o cluster tentava aumentar ou reduzir a escala verticalmente.
- Corrigido problema com falhas de etapas ou tarefas durante a escalabilidade do cluster ao ser garantido que os estados dos nós fossem sempre consistentes entre os daemons do Amazon EMR no cluster e o YARN/HDFS.
- Corrigido um problema em que operações de cluster, como redução de escala verticalmente e envio de etapas, falhavam para clusters do Amazon EMR habilitados com a autenticação Kerberos. Isso ocorreu porque o daemon no cluster do Amazon EMR não renovou o tíquete Kerberos, que é necessário para se comunicar com segurança com a execução no nó primário. HDFS/YARN
- Configurar um cluster para corrigir problemas de desempenho do servidor de linha do tempo do Apache YARN versões 1 e 1.5

  As versões 1 e 1.5 do servidor de linha do tempo do Apache YARN podem causar problemas de performance com clusters do EMR muito ativos e grandes, especialmente com `yarn.resourcemanager.system-metrics-publisher.enabled=true`, que é a configuração padrão no Amazon EMR. Um servidor de linha do tempo do YARN v2 de código

aberto resolve o problema de performance relacionado à escalabilidade do servidor de linha do tempo do YARN.

Outras soluções alternativas para esse problema incluem:

- Configurando yarn.resourcemanager. system-metrics-publisher.enabled=false em yarn-site.xml.
- Habilitar a correção para esse problema na criação de um cluster, conforme descrito abaixo.

As seguintes versões do Amazon EMR contêm uma correção para esse problema de performance do servidor de linha do tempo do YARN.

EMR 5.30.2, 5.31.1, 5.32.1, 5.33.1, 5.34.x, 6.0.1, 6.1.1, 6.2.1, 6.3.1, 6.4.x

Para habilitar a correção em qualquer uma das versões do Amazon EMR especificadas acima, defina essas propriedades como `true` em um arquivo JSON de configurações que é passado usando o parâmetro de comando `aws emr create-cluster`: `--configurations file://./configurations.json`. Ou habilite a correção usando a interface do usuário do console de reconfiguração.

Exemplo de conteúdo do arquivo configurations.json:

```
[
{
"Classification": "yarn-site",
"Properties": {
"yarn.resourcemanager.system-metrics-publisher.timeline-server-v1.enable-batch":
 "true",
"yarn.resourcemanager.system-metrics-publisher.enabled": "true"
},
"Configurations": []
}
]
```

- WebHDFS e o servidor HttpFS estão desabilitados por padrão. Você pode reabilitar o WebHDFS usando a configuração do Hadoop, `dfs.webhdfs.enabled`. O servidor HttpFS pode ser iniciado usando `sudo systemctl start hadoop-httpfs`.
- O HTTPS agora está habilitado por padrão para repositórios do Amazon Linux. Se você estiver usando uma política de VPCE do Amazon S3 para restringir o acesso a buckets específicos, deverá adicionar o novo ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` do bucket do Amazon Linux à sua política (substitua `$region` pela região em que o endpoint está situado). Para obter mais informações, consulte esse tópico nos fóruns de AWS discussão. Anúncio: o

Amazon Linux 2 agora oferece suporte à capacidade de usar HTTPS ao se conectar a repositórios de pacotes .

- Hive: a performance da consulta de gravação foi aprimorada ao ser permitido o uso de um diretório temporário no HDFS para o último trabalho. Os dados temporários do trabalho final são gravados no HDFS em vez de no Amazon S3 e a performance é melhorada porque os dados são movidos do HDFS para a localização da tabela final (Amazon S3) em vez de entre dispositivos do Amazon S3.
- Hive: melhoria do tempo de compilação de consultas em até 2,5 vezes com a remoção de partições de metastores do Glue.
- Por padrão, quando UDFs os integrados são passados pelo Hive para o Hive Metastore Server, somente um subconjunto desses incorporados é UDFs passado para o Glue Metastore, já que o Glue suporta apenas operadores de expressão limitados. Se você definir `hive.glue.partition.pruning.client=true`, toda a remoção de partições ocorrerá no lado do cliente. Se você definir `hive.glue.partition.pruning.server=true`, toda a remoção de partições ocorrerá no lado do servidor.

Problemas conhecidos

- As consultas no Hue não funcionam no Amazon EMR 6.4.0 porque o servidor HttpFS do Apache Hadoop está desabilitado por padrão. Para usar o Hue no Amazon EMR 6.4.0, inicie manualmente o servidor HttpFS no nó primário do Amazon EMR usando `sudo systemctl start hadoop-httpfs`, ou use uma etapa do Amazon EMR.
- O atributo Cadernos do Amazon EMR usado com a personificação de usuários do Livy não funciona porque o HttpFS está desabilitado por padrão. Nesse caso, o caderno do EMR não pode se conectar ao cluster que tem a personificação do Livy habilitada. A solução alternativa é iniciar o servidor HttpFS antes de conectar o caderno do EMR ao cluster usando `sudo systemctl start hadoop-httpfs`.
- Na versão 6.4.0 do Amazon EMR, o Phoenix não é compatível com o componente de conectores do Phoenix.
- Para usar as ações do Spark com o Apache Oozie, você deve adicionar a seguinte configuração ao seu arquivo `workflow.xml` do Oozie. Caso contrário, várias bibliotecas críticas, como Hadoop e EMRFS, estarão ausentes do classpath dos executores do Spark que o Oozie inicia.

```
<spark-opts>--conf spark.yarn.populateHadoopClasspath=true</spark-opts>
```

- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
  - O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

  Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.

Versão 5.32.0

As notas da versão a seguir incluem informações para a versão 5.32.0 do Amazon EMR. As alterações são referentes à versão 5.31.0.

Data da versão inicial: 8 de janeiro de 2021

Atualizações

- Atualizado o conector do Amazon Glue para a versão 1.14.0
- SDK do Amazon SageMaker Spark atualizado para a versão 1.4.1
- Atualizado para AWS SDK para Java a versão 1.11.890
- Atualizado o conector do DynamoDB para EMR, versão 4.16.0
- Atualizado o EMRFS para a versão 2.45.0
- Atualizadas as métricas de analytics de log do EMR para a versão 1.18.0
- Cliente MetricsAndEventsApiGateway EMR atualizado para a versão 1.5.0
- Atualizado o servidor de registros do EMR para a versão 1.8.0

- Atualizado o EMR S3 Dist CP para a versão 2.17.0
- Atualizado o agente secreto do EMR para a versão 1.7.0
- Atualizado o Flink para a versão 1.11.2
- Atualizado o Hadoop para a versão 2.10.1-amzn-0
- Atualizado o Hive para a versão 2.3.7-amzn-3
- Atualizado o Hue para a versão 4.8.0
- Atualizado o Mxnet para a versão 1.7.0
- Atualizado o OpenCV para a versão 4.4.0
- Atualizado o Presto para a versão 0.240.1-amzn-0
- Atualizado o Spark para a versão 2.4.7-amnz-0
- Atualizado TensorFlow para a versão 2.3.1

Alterações, melhorias e problemas resolvidos

- Esta é uma versão para corrigir problemas com o Amazon EMR Scaling quando ele não consegue up/scale reduzir a escala de um cluster com sucesso ou causa falhas no aplicativo.
- Corrigido um problema em que as solicitações de escalabilidade falhavam em um cluster grande e altamente utilizado quando os daemons do Amazon EMR no cluster estavam executando atividades de verificação de integridade, como a coleta do estado do nó do YARN e o estado do nó do HDFS. Isso estava acontecendo porque os daemons no cluster não conseguiam comunicar os dados do status de integridade de um nó aos componentes internos do Amazon EMR.
- Aprimorados os daemons do EMR no cluster para rastrear corretamente os estados dos nós quando são reutilizados endereços IP para melhorar a confiabilidade durante operações de escalabilidade.
- SPARK-29683. Corrigido um problema em que ocorriam falhas de trabalho durante a redução da escala verticalmente do cluster, pois o Spark presumia que todos os nós disponíveis estavam na lista de negação.
- YARN-9011. Corrigido um problema em que ocorriam falhas de trabalho devido a uma condição de corrida na desativação do YARN quando o cluster tentava aumentar ou reduzir a escala verticalmente.
- Corrigido problema com falhas de etapas ou tarefas durante a escalabilidade do cluster ao ser garantido que os estados dos nós fossem sempre consistentes entre os daemons do Amazon EMR no cluster e o YARN/HDFS.

- Corrigido um problema em que operações de cluster, como redução de escala verticalmente e envio de etapas, falhavam para clusters do Amazon EMR habilitados com a autenticação Kerberos. Isso ocorreu porque o daemon no cluster do Amazon EMR não renovou o tíquete Kerberos, que é necessário para se comunicar com segurança com a execução no nó primário. HDFS/YARN
- As versões mais recentes do Amazon EMR corrigem o problema com um limite menor de “Máximo de arquivos abertos” em relação às versões mais antigas AL2 no Amazon EMR. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR agora incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”.
- Atualizadas versões do componente.
- Para obter uma lista das versões do componente, consulte Sobre versões do Amazon EMR neste guia.

Novos recursos

- Desde as versões 5.32.0 e 6.5.0 do Amazon EMR, o dimensionamento do executor dinâmico para o Apache Spark está habilitado por padrão. Para ativar ou desativar esse atributo, você pode usar o parâmetro de configuração `spark.yarn.heterogeneousExecutors.enabled`.
- Status de suporte do Instance Metadata Service (IMDS) V2: os componentes do Amazon EMR 5.23.1, 5.27.1 e 5.32 ou posteriores são usados para todas as chamadas do IMDS. IMDSv2 Para chamadas IMDS no código do aplicativo, você pode usar ambos IMDSv1 e IMDSv2, ou configurar o IMDS para uso somente IMDSv2 para aumentar a segurança. Para outras versões 5.x do EMR, a IMDSv1 desativação causa falha na inicialização do cluster.
- Desde a versão 5.32.0 do Amazon EMR, você pode iniciar um cluster que se integre nativamente ao Apache Ranger. O Apache Ranger é uma estrutura de código aberto para habilitar, monitorar e gerenciar uma segurança de dados abrangente em toda a plataforma Hadoop. Para obter mais informações, consulte Apache Ranger. Com a integração nativa, você pode trazer seu próprio Apache Ranger para aplicar um controle de acesso detalhado aos dados no Amazon EMR. Consulte Integrar o Amazon EMR com o Apache Ranger no Guia de lançamento do Amazon EMR.
- A versão 5.32.0 do Amazon EMR é compatível com o Amazon EMR no EKS. Para obter mais detalhes sobre os conceitos básicos do EMR no EKS, consulte O que é o Amazon EMR no EKS.
- A versão 5.32.0 do Amazon EMR é compatível com o Amazon EMR Studio (versão de demonstração). Para obter mais detalhes sobre os conceitos básicos do EMR Studio, consulte Amazon EMR Studio (versão de demonstração).

- Políticas gerenciadas com escopo definido: para se alinhar às AWS melhores práticas, o Amazon EMR introduziu políticas gerenciadas padrão com escopo do EMR v2 como substitutas das políticas que serão descontinuadas. Consulte as políticas gerenciadas do Amazon EMR.

Problemas conhecidos

- Para clusters de sub-rede privados do Amazon EMR 6.3.0 e 6.2.0, você não pode acessar a interface do usuário da Web do Ganglia. Você receberá um erro de “acesso negado (403)”. Outros sites UIs, como Spark, Hue, Zeppelin JupyterHub, Livy e Tez, estão funcionando normalmente. O acesso à interface do usuário da Web do Ganglia em clusters de sub-redes públicas também está funcionando normalmente. Para resolver esse problema, reinicie o serviço httpd no nó primário com `sudo systemctl restart httpd`. Esse problema foi corrigido na versão 6.4.0 do Amazon EMR.
- Limite inferior de “Máximo de arquivos abertos” em versões mais antigas AL2 [corrigido em versões mais recentes]. Versões do Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 e emr-6.2.0 são baseadas em versões mais antigas do Amazon Linux 2 (), que AL2 têm uma configuração de limite inferior para “Máximo de arquivos abertos” quando clusters do Amazon EMR são criados com a AMI padrão. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”. Versões com o limite inferior de arquivos abertos causam o erro “Muitos arquivos abertos” ao ser enviado um trabalho do Spark. Nas versões afetadas, a AMI padrão do Amazon EMR tem uma configuração de ulimit padrão de 4096 para “Máximo de arquivos abertos”, que é inferior ao limite de 65536 arquivos na AMI mais recente do Amazon Linux 2. A configuração inferior de ulimit para “Máximo de arquivos abertos” causa falhas em trabalhos do Spark quando o driver e o executor do Spark tentam abrir mais de 4096 arquivos. Para corrigir o problema, o Amazon EMR tem um script de ação de bootstrap (BA) que ajusta a configuração de ulimit na criação do cluster.

  Se você está usando uma versão mais antiga do Amazon EMR que não tem a correção permanente para esse problema, a solução alternativa a seguir permite que você defina explicitamente o ulimit instance-controller para um máximo de 65536 arquivos.

  Defina explicitamente um ulimit na linha de comando

  1. Edite `/etc/systemd/system/instance-controller.service` para adicionar os seguintes parâmetros à seção Serviço.

     ```
     LimitNOFILE=65536
     ```

```
LimitNPROC=65536
```

2. Reiniciar InstanceController

```
$ sudo systemctl daemon-reload

$ sudo systemctl restart instance-controller
```

Defina um ulimit usando a ação de bootstrap (BA)

Você também pode usar um script de ação de bootstrap (BA) para configurar o ulimit instance-controller para 65536 arquivos na criação do cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- **Important**

  Os clusters do EMR que executam Amazon Linux ou Amazon Linux 2 Amazon Machine Images (AMIs) usam o comportamento padrão do Amazon Linux e não baixam e instalam automaticamente atualizações importantes e críticas do kernel que exigem uma reinicialização. Esse é o mesmo comportamento de outras EC2 instâncias da Amazon que executam o Amazon Linux AMI padrão. Se novas atualizações de software do Amazon Linux que exigem reinicialização (como atualizações do kernel, NVIDIA e CUDA) forem disponibilizadas após o lançamento de uma versão do Amazon EMR, as instâncias de

cluster do Amazon EMR que executam a AMI padrão não baixarão nem instalarão essas atualizações automaticamente. Para obter atualizações do kernel, você pode personalizar sua AMI do Amazon EMR para usar a AMI do Amazon Linux mais recente.

- O suporte do console para criar uma configuração de segurança que especifica a opção de integração do AWS Ranger atualmente não é suportado na GovCloud região. A configuração de segurança do pode ser feita usando a CLI. Consulte Criar a configuração de segurança do EMR no Guia de gerenciamento do Amazon EMR.
- Quando AtRestEncryption a criptografia HDFS é habilitada em um cluster que usa o Amazon EMR 5.31.0 ou 5.32.0, as consultas do Hive resultam na seguinte exceção de tempo de execução.

```
TaskAttempt 3 failed, info=[Error: Error while running task ( failure ) :
 attempt_1604112648850_0001_1_01_000000_3:java.lang.RuntimeException:
 java.lang.RuntimeException: Hive Runtime Error while closing
 operators: java.io.IOException: java.util.ServiceConfigurationError:
 org.apache.hadoop.security.token.TokenIdentifier: Provider
 org.apache.hadoop.hbase.security.token.AuthenticationTokenIdentifier not found
```

- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
  - O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

  Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.

Versão 6.2.0

As notas da versão a seguir incluem informações para a versão 6.2.0 do Amazon EMR. As alterações são referentes à versão 6.1.0.

Data da versão inicial: 9 de dezembro de 2020

Data da última atualização: 4 de outubro de 2021

Aplicações compatíveis

- AWS SDK para Java versão 1.11.828
- emr-record-server versão 1.7.0
- Flink versão 1.11.2
- Ganglia versão 3.7.2
- Hadoop versão 3.2.1-amzn-1
- HBase versão 2.2.6-amzn-0
- HBase-operator-tools 1.0.0
- HCatalog versão 3.1.2-amzn-0
- Hive versão 3.1.2-amzn-3
- Hudi versão 0.6.0-amzn-1
- Hue versão 4.8.0
- JupyterHub versão 1.1.0
- Livy versão 0.7.0
- MXNet versão 1.7.0
- Oozie versão 5.2.0
- Phoenix versão 5.0.0
- Pig versão 0.17.0
- Presto versão 0.238.3-amzn-1
- PrestoSQL versão 343
- Spark versão 3.0.1-amzn-0
- spark-rapids 0.2.0
- TensorFlow versão 2.3.1
- Zeppelin versão 0.9.0-preview1

- Zookeeper versão 3.4.14
- Conectores e drivers: DynamoDB Connector 4.16.0

Novos recursos

- HBase: a renomeação foi removida na fase de confirmação e o HFile rastreamento persistente foi adicionado. Consulte HFile Rastreamento persistente no Guia de lançamento do Amazon EMR.
- HBase: Backported Crie uma configuração que força o armazenamento em cache dos blocos na compactação.
- PrestoDB: melhorias na remoção dinâmica de partições. O Join Reorder baseado em regras funciona em dados não particionados.
- Políticas gerenciadas com escopo definido: para se alinhar às AWS melhores práticas, o Amazon EMR introduziu políticas gerenciadas padrão com escopo do EMR v2 como substitutas das políticas que serão descontinuadas. Consulte as políticas gerenciadas do Amazon EMR.
- Status de suporte do Instance Metadata Service (IMDS) V2: Para o Amazon EMR 6.2 ou posterior, os componentes do Amazon EMR são usados para todas as chamadas do IMDS. IMDSv2 Para chamadas IMDS no código do aplicativo, você pode usar ambos IMDSv1 e IMDSv2, ou configurar o IMDS para uso somente IMDSv2 para aumentar a segurança. Se você desabilitar IMDSv1 em versões anteriores do Amazon EMR 6.x, isso causará uma falha na inicialização do cluster.

Alterações, melhorias e problemas resolvidos

- Esta é uma versão para corrigir problemas com o Amazon EMR Scaling quando ele não consegue up/scale reduzir a escala de um cluster com sucesso ou causa falhas no aplicativo.
- Corrigido um problema em que as solicitações de escalabilidade falhavam em um cluster grande e altamente utilizado quando os daemons do Amazon EMR no cluster estavam executando atividades de verificação de integridade, como a coleta do estado do nó do YARN e o estado do nó do HDFS. Isso estava acontecendo porque os daemons no cluster não conseguiam comunicar os dados do status de integridade de um nó aos componentes internos do Amazon EMR.
- Aprimorados os daemons do EMR no cluster para rastrear corretamente os estados dos nós quando são reutilizados endereços IP para melhorar a confiabilidade durante operações de escalabilidade.
- SPARK-29683. Corrigido um problema em que ocorriam falhas de trabalho durante a redução da escala verticalmente do cluster, pois o Spark presumia que todos os nós disponíveis estavam na lista de negação.

- YARN-9011. Corrigido um problema em que ocorriam falhas de trabalho devido a uma condição de corrida na desativação do YARN quando o cluster tentava aumentar ou reduzir a escala verticalmente.
- Corrigido problema com falhas de etapas ou tarefas durante a escalabilidade do cluster ao ser garantido que os estados dos nós fossem sempre consistentes entre os daemons do Amazon EMR no cluster e o YARN/HDFS.
- Corrigido um problema em que operações de cluster, como redução de escala verticalmente e envio de etapas, falhavam para clusters do Amazon EMR habilitados com a autenticação Kerberos. Isso ocorreu porque o daemon no cluster do Amazon EMR não renovou o tíquete Kerberos, que é necessário para se comunicar com segurança com a execução no nó primário. HDFS/YARN
- As versões mais recentes do Amazon EMR corrigem o problema com um limite menor de “Máximo de arquivos abertos” em relação às versões mais antigas AL2 no Amazon EMR. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR agora incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”.
- Spark: melhorias de performance no runtime do Spark.

Problemas conhecidos

- O Amazon EMR 6.2 tem permissões incorretas definidas no diretório/etc/cron.d/libinstance-controller-java file in EMR 6.2.0. Permissions on the file are 645 (-rw-r--r-x), when they should be 644 (-rw-r--r--). As a result, Amazon EMR version 6.2 does not log instance-state logs, and the /emr/instance-logs vazio. Esse problema foi corrigido nas versões 6.3.0 e posteriores do Amazon EMR.

  Para contornar esse problema, execute o script a seguir como uma ação de bootstrap na inicialização do cluster.

  ```
  #!/bin/bash
  sudo chmod 644 /etc/cron.d/libinstance-controller-java
  ```

- Para clusters de sub-rede privados do Amazon EMR 6.2.0 e 6.3.0, você não pode acessar a interface do usuário da Web do Ganglia. Você receberá um erro de “acesso negado (403)”. Outros sites UIs, como Spark, Hue, Zeppelin JupyterHub, Livy e Tez, estão funcionando normalmente. O acesso à interface do usuário da Web do Ganglia em clusters de sub-redes públicas também

está funcionando normalmente. Para resolver esse problema, reinicie o serviço httpd no nó primário com `sudo systemctl restart httpd`. Esse problema foi corrigido na versão 6.4.0 do Amazon EMR.

- Há um problema no Amazon EMR 6.2.0 em que o httpd falha continuamente, fazendo com que o Ganglia fique indisponível. Você recebe a mensagem de erro “cannot connect to the server”. Para corrigir um cluster que já está em execução com esse problema, use SSH no nó primário do cluster e adicione a linha `Listen 80` ao arquivo `httpd.conf` localizado em `/etc/httpd/conf/httpd.conf`. Esse problema foi corrigido na versão 6.3.0 do Amazon EMR.
- O HTTPD apresenta falha nos clusters do EMR 6.2.0 quando é usada uma configuração de segurança. Isso faz com que a interface de usuário da aplicação Web do Ganglia fique indisponível. Para acessar a interface de usuário da aplicação Web do Ganglia, adicione `Listen 80` ao arquivo `/etc/httpd/conf/httpd.conf` no nó primário do cluster. Para obter mais informações sobre como se conectar ao cluster, consulte Connect to the Primary Node Using SSH.

  Os Cadernos do EMR também não conseguem estabelecer uma conexão com clusters do EMR 6.2.0 quando você usa uma configuração de segurança. O caderno não conseguirá listar os kernels e enviar trabalhos do Spark. Em vez disso, recomendamos que você use os Cadernos do EMR com outra versão do Amazon EMR.
- Limite inferior de “Máximo de arquivos abertos” em versões mais antigas AL2 [corrigido em versões mais recentes]. Versões do Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 e emr-6.2.0 são baseadas em versões mais antigas do Amazon Linux 2 (), que AL2 têm uma configuração de limite inferior para “Máximo de arquivos abertos” quando clusters do Amazon EMR são criados com a AMI padrão. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”. Versões com o limite inferior de arquivos abertos causam o erro “Muitos arquivos abertos” ao ser enviado um trabalho do Spark. Nas versões afetadas, a AMI padrão do Amazon EMR tem uma configuração de ulimit padrão de 4096 para “Máximo de arquivos abertos”, que é inferior ao limite de 65536 arquivos na AMI mais recente do Amazon Linux 2. A configuração inferior de ulimit para “Máximo de arquivos abertos” causa falhas em trabalhos do Spark quando o driver e o executor do Spark tentam abrir mais de 4096 arquivos. Para corrigir o problema, o Amazon EMR tem um script de ação de bootstrap (BA) que ajusta a configuração de ulimit na criação do cluster.

  Se você está usando uma versão mais antiga do Amazon EMR que não tem a correção permanente para esse problema, a solução alternativa a seguir permite que você defina explicitamente o ulimit instance-controller para um máximo de 65536 arquivos.

Defina explicitamente um ulimit na linha de comando

1. Edite `/etc/systemd/system/instance-controller.service` para adicionar os seguintes parâmetros à seção Serviço.

   ```
   LimitNOFILE=65536
   ```

   ```
   LimitNPROC=65536
   ```

2. Reiniciar InstanceController

   ```
   $ sudo systemctl daemon-reload
   ```

   ```
   $ sudo systemctl restart instance-controller
   ```

Defina um ulimit usando a ação de bootstrap (BA)

Você também pode usar um script de ação de bootstrap (BA) para configurar o ulimit instance-controller para 65536 arquivos na criação do cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- > **Important**
  >
  > O Amazon EMR 6.1.0 e 6.2.0 incluem um problema de performance que pode afetar criticamente todas as operações de inserção, upsert e exclusão do Hudi. Se você planeja usar o Hudi com o Amazon EMR 6.1.0 ou 6.2.0, entre em AWS contato com o suporte para obter um Hudi RPM corrigido.

- > **Important**
  >
  > Os clusters do EMR que executam Amazon Linux ou Amazon Linux 2 Amazon Machine Images (AMIs) usam o comportamento padrão do Amazon Linux e não baixam e instalam automaticamente atualizações importantes e críticas do kernel que exigem uma reinicialização. Esse é o mesmo comportamento de outras EC2 instâncias da Amazon que executam o Amazon Linux AMI padrão. Se novas atualizações de software do Amazon Linux que exigem reinicialização (como atualizações do kernel, NVIDIA e CUDA) forem disponibilizadas após o lançamento de uma versão do Amazon EMR, as instâncias de cluster do Amazon EMR que executam a AMI padrão não baixarão nem instalarão essas atualizações automaticamente. Para obter atualizações do kernel, você pode personalizar sua AMI do Amazon EMR para usar a AMI do Amazon Linux mais recente.

- Os artefatos do Maven do Amazon EMR 6.2.0 não são publicados. Eles serão publicados com uma versão futura do Amazon EMR.
- O HFile rastreamento persistente usando a tabela do sistema HBase storefile não oferece suporte ao recurso de replicação da HBase região. Para obter mais informações sobre a replicação HBase da região, consulte Leituras de alta disponibilidade consistentes com o cronograma.
- Diferenças entre as versões de bucketing do Amazon EMR 6.x e do EMR 5.x Hive

  O EMR 5.x usa o OOS Apache Hive 2, enquanto o EMR 6.x usa o OOS Apache Hive 3. O Hive2 de código aberto usa o Bucketing versão 1, enquanto o Hive3 de código aberto usa o Bucketing versão 2. Essa diferença de versão de bucketing entre o Hive 2 (EMR 5.x) e o Hive 3 (EMR 6.x) significa que o hash de bucketing do Hive funciona de uma forma diferente. Veja o exemplo abaixo.

  A tabela a seguir é um exemplo criado no EMR 6.x e no EMR 5.x, respectivamente.

```
-- Using following LOCATION in EMR 6.x
CREATE TABLE test_bucketing (id INT, desc STRING)
```

```
PARTITIONED BY (day STRING)
CLUSTERED BY(id) INTO 128 BUCKETS
LOCATION 's3://your-own-s3-bucket/emr-6-bucketing/';

-- Using following LOCATION in EMR 5.x
LOCATION 's3://your-own-s3-bucket/emr-5-bucketing/';
```

Inserir os mesmos dados no EMR 6.x e no EMR 5.x.

```
INSERT INTO test_bucketing PARTITION (day='01') VALUES(66, 'some_data');
INSERT INTO test_bucketing PARTITION (day='01') VALUES(200, 'some_data');
```

A verificação da localização do S3 mostra que o nome do arquivo de bucketing é diferente, pois a função de hash é diferente entre o EMR 6.x (Hive 3) e o EMR 5.x (Hive 2).

```
[hadoop@ip-10-0-0-122 ~]$ aws s3 ls s3://your-own-s3-bucket/emr-6-bucketing/day=01/
2020-10-21 20:35:16         13 000025_0
2020-10-21 20:35:22         14 000121_0
[hadoop@ip-10-0-0-122 ~]$ aws s3 ls s3://your-own-s3-bucket/emr-5-bucketing/day=01/
2020-10-21 20:32:07         13 000066_0
2020-10-21 20:32:51         14 000072_0
```

Você também pode ver a diferença de versão executando o comando a seguir na CLI do Hive no EMR 6.x. Observe que ele retorna a versão 2 do bucketing.

```
hive> DESCRIBE FORMATTED test_bucketing;
...
Table Parameters:
    bucketing_version       2
...
```

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

Solução:

- SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
- Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
  - O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

  Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.

Versão 5.31.0

As notas da versão a seguir incluem informações para a versão 5.31.0 do Amazon EMR. As alterações são referentes à versão 5.30.1.

Data da versão inicial: 9 de outubro de 2020

Data da última atualização: 15 de outubro de 2020

Atualizações

- Atualizado o conector do Amazon Glue para a versão 1.13.0
- SDK do Amazon SageMaker Spark atualizado para a versão 1.4.0
- Atualizado o conector do Amazon Kinesis para a versão 3.5.9
- Atualizado para AWS SDK para Java a versão 1.11.852
- Atualizado o Bigtop-tomcat para a versão 8.5.56
- Atualizado o EMRFS para a versão 2.43.0
- Cliente MetricsAndEventsApiGateway EMR atualizado para a versão 1.4.0
- Atualizado o EMR S3 Dist CP para a versão 2.15.0
- Atualizado o EMR S3 Select para a versão 1.6.0
- Atualizado o Flink para a versão 1.11.0
- Atualizado o Hadoop para a versão 2.10.0
- Atualizado o Hive para a versão 2.3.7
- Atualizado o Hudi para a versão 0.6.0
- Atualizado o Hue para a versão 4.7.1
- Atualizado JupyterHub para a versão 1.1.0
- Atualizado o Mxnet para a versão 1.6.0
- Atualizado o OpenCV para a versão 4.3.0
- Atualizado o Presto para a versão 0.238.3
- Atualizado TensorFlow para a versão 2.1.0

Alterações, melhorias e problemas resolvidos

- Esta é uma versão para corrigir problemas com o Amazon EMR Scaling quando ele não consegue up/scale reduzir a escala de um cluster com sucesso ou causa falhas no aplicativo.

- Corrigido um problema em que as solicitações de escalabilidade falhavam em um cluster grande e altamente utilizado quando os daemons do Amazon EMR no cluster estavam executando atividades de verificação de integridade, como a coleta do estado do nó do YARN e o estado do nó do HDFS. Isso estava acontecendo porque os daemons no cluster não conseguiam comunicar os dados do status de integridade de um nó aos componentes internos do Amazon EMR.
- Aprimorados os daemons do EMR no cluster para rastrear corretamente os estados dos nós quando são reutilizados endereços IP para melhorar a confiabilidade durante operações de escalabilidade.
- SPARK-29683. Corrigido um problema em que ocorriam falhas de trabalho durante a redução da escala verticalmente do cluster, pois o Spark presumia que todos os nós disponíveis estavam na lista de negação.
- YARN-9011. Corrigido um problema em que ocorriam falhas de trabalho devido a uma condição de corrida na desativação do YARN quando o cluster tentava aumentar ou reduzir a escala verticalmente.
- Corrigido problema com falhas de etapas ou tarefas durante a escalabilidade do cluster ao ser garantido que os estados dos nós fossem sempre consistentes entre os daemons do Amazon EMR no cluster e o YARN/HDFS.
- Corrigido um problema em que operações de cluster, como redução de escala verticalmente e envio de etapas, falhavam para clusters do Amazon EMR habilitados com a autenticação Kerberos. Isso ocorreu porque o daemon no cluster do Amazon EMR não renovou o tíquete Kerberos, que é necessário para se comunicar com segurança com a execução no nó primário. HDFS/YARN
- As versões mais recentes do Amazon EMR corrigem o problema com um limite menor de “Máximo de arquivos abertos” em relação às versões mais antigas AL2 no Amazon EMR. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR agora incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”.
- As estatísticas de colunas do Hive são compatíveis com as versões 5.31.0 e posteriores do Amazon EMR.
- Atualizadas versões do componente.
- Suporte ao EMRFS S3EC V2 no Amazon EMR 5.31.0. Nas versões 1.11.837 e posteriores do SDK para Java no S3, a versão 2 do cliente de criptografia (S3EC V2) foi introduzida com vários aprimoramentos de segurança. Para obter mais informações, consulte:
  - Publicações no blog do S3: Atualizações no cliente de criptografia do Amazon S3.

- AWS SDK para Java Guia do desenvolvedor: Migre clientes de criptografia e descriptografia para a V2.
- Guia de gerenciamento do EMR: Criptografia do lado do cliente do Amazon S3.

O Encryption Client V1 ainda está disponível no SDK para compatibilidade retroativa.

Novos recursos

- Limite inferior de “Máximo de arquivos abertos” em versões mais antigas AL2 [corrigido em versões mais recentes]. Versões do Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 e emr-6.2.0 são baseadas em versões mais antigas do Amazon Linux 2 (), que AL2 têm uma configuração de limite inferior para “Máximo de arquivos abertos” quando clusters do Amazon EMR são criados com a AMI padrão. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”. Versões com o limite inferior de arquivos abertos causam o erro “Muitos arquivos abertos” ao ser enviado um trabalho do Spark. Nas versões afetadas, a AMI padrão do Amazon EMR tem uma configuração de ulimit padrão de 4096 para “Máximo de arquivos abertos”, que é inferior ao limite de 65536 arquivos na AMI mais recente do Amazon Linux 2. A configuração inferior de ulimit para “Máximo de arquivos abertos” causa falhas em trabalhos do Spark quando o driver e o executor do Spark tentam abrir mais de 4096 arquivos. Para corrigir o problema, o Amazon EMR tem um script de ação de bootstrap (BA) que ajusta a configuração de ulimit na criação do cluster.

  Se você está usando uma versão mais antiga do Amazon EMR que não tem a correção permanente para esse problema, a solução alternativa a seguir permite que você defina explicitamente o ulimit instance-controller para um máximo de 65536 arquivos.

  Defina explicitamente um ulimit na linha de comando

  1. Edite `/etc/systemd/system/instance-controller.service` para adicionar os seguintes parâmetros à seção Serviço.

     ```
     LimitNOFILE=65536

     LimitNPROC=65536
     ```

  2. Reiniciar InstanceController

     ```
     $ sudo systemctl daemon-reload
     ```

```
$ sudo systemctl restart instance-controller
```

Defina um ulimit usando a ação de bootstrap (BA)

Você também pode usar um script de ação de bootstrap (BA) para configurar o ulimit instance-controller para 65536 arquivos na criação do cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- Com a versão 5.31.0 do Amazon EMR, você pode iniciar um cluster que se integre ao Lake Formation. Essa integração fornece filtragem de dados refinada em nível de coluna para bancos de dados e tabelas no Glue Data Catalog. AWS Ela também permite logon único federado para Cadernos do EMR ou para o Apache Zeppelin em um sistema de identidade empresarial. Para obter mais informações, consulte Integrating Amazon EMR with AWS Lake Formation no Guia de gerenciamento do Amazon EMR.

  Atualmente, o Amazon EMR com Lake Formation está disponível em 16 AWS regiões: Leste dos EUA (Ohio e Norte da Virgínia), Oeste dos EUA (Norte da Califórnia e Oregon), Ásia-Pacífico (Mumbai, Seul, Cingapura, Sydney e Tóquio), Canadá (Central), Europa (Frankfurt, Irlanda, Londres, Paris e Estocolmo) e América do Sul (São Paulo).

Problemas conhecidos

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

  Solução:

  - SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
  - Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

- Quando AtRestEncryption a criptografia HDFS é habilitada em um cluster que usa o Amazon EMR 5.31.0 ou 5.32.0, as consultas do Hive resultam na seguinte exceção de tempo de execução.

  ```
  TaskAttempt 3 failed, info=[Error: Error while running task ( failure ) :
   attempt_1604112648850_0001_1_01_000000_3:java.lang.RuntimeException:
   java.lang.RuntimeException: Hive Runtime Error while closing
   operators: java.io.IOException: java.util.ServiceConfigurationError:
   org.apache.hadoop.security.token.TokenIdentifier: Provider
   org.apache.hadoop.hbase.security.token.AuthenticationTokenIdentifier not found
  ```

- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
  - O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

  Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.

Versão 6.1.0

As notas da versão a seguir incluem informações para a versão 6.1.0 do Amazon EMR. As alterações são referentes à versão 6.0.0.

Data da versão inicial: 4 de setembro de 2020

Data da última atualização: 15 de outubro de 2020

Aplicações compatíveis

- AWS SDK para Java versão 1.11.828
- Flink versão 1.11.0
- Ganglia versão 3.7.2
- Hadoop versão 3.2.1-amzn-1
- HBase versão 2.2.5
- HBase-operator-tools 1.0.0
- HCatalog versão 3.1.2-amzn-0

- Hive versão 3.1.2-amzn-1
- Hudi versão 0.5.2-incubating
- Hue versão 4.7.1
- JupyterHub versão 1.1.0
- Livy versão 0.7.0
- MXNet versão 1.6.0
- Oozie versão 5.2.0
- Phoenix versão 5.0.0
- Presto versão 0.232
- PrestoSQL versão 338
- Spark versão 3.0.0-amzn-0
- TensorFlow versão 2.1.0
- Zeppelin versão 0.9.0-preview1
- Zookeeper versão 3.4.14
- Conectores e drivers: DynamoDB Connector 4.14.0

Novos recursos

- Os tipos de instância ARM são compatíveis desde o Amazon EMR versão 5.30.0 e do Amazon EMR versão 6.1.0.
- Os tipos de instância de uso geral M6g são compatíveis desde as versões 6.1.0 e 5.30.0 do Amazon EMR. Para obter mais informações, consulte Tipos de instâncias compatíveis no Guia de gerenciamento do Amazon EMR.
- O recurso de grupos de EC2 posicionamento é suportado a partir da versão 5.23.0 do Amazon EMR como uma opção para vários clusters de nós primários. Atualmente, somente os tipos de nós primários são compatíveis com o atributo grupo de posicionamento e a estratégia `SPREAD` é aplicada a estes nós primários. A estratégia `SPREAD` posiciona um pequeno grupo de instâncias em um hardware subjacente separado para evitar a perda de múltiplos nós primários em caso de falha de hardware. Para obter mais informações, consulte Integração do EMR com o EC2 Placement Group no Guia de gerenciamento do Amazon EMR.
- Ajuste de escala gerenciado: com a versão 6.1.0 do Amazon EMR, é possível habilitar o Ajuste de Escala Gerenciado do Amazon EMR para aumentar ou diminuir automaticamente o número de instâncias ou unidades no cluster com base na workload. O Amazon EMR avalia continuamente as

métricas do cluster para tomar decisões de ajuste de escala que otimizam os clusters em termos de custo e velocidade. O ajuste de escala gerenciado também está disponível na versão 5.30.0 e posteriores do Amazon EMR, exceto na versão 6.0.0. Para obter mais informações, consulte Escalar recursos de cluster no Guia de gerenciamento do Amazon EMR.

- O PrestoSQL versão 338 é compatível com o EMR 6.1.0. Para obter mais informações, consulte Presto.
  - O PrestoSQL só é compatível com as versões 6.1.0 e posteriores do EMR e não no EMR 6.0.0 ou no EMR 5.x.
  - O nome da aplicação, `Presto`, continua a ser usado para instalar o PrestoDB em clusters. Para instalar o PrestoSQL em clusters, use o nome da aplicação `PrestoSQL`.
  - Você pode instalar o PrestoDB ou o PrestoSQL, mas não pode instalar os dois em um único cluster. Se o PrestoDB e o PrestoSQL forem especificados na tentativa de criação de um cluster, ocorrerá um erro de validação e a solicitação de criação do cluster falhará.
  - O PrestoSQL é compatível com clusters de mestre único e de vários mestres. Em clusters de vários mestres, é necessário um metastore externo do Hive para executar o PrestoSQL ou o PrestoDB. Consulte Aplicações compatíveis em um cluster do EMR com vários nós primários.
- Suporte à autenticação automática do ECR no Apache Hadoop e no Apache Spark com Docker: os usuários do Spark podem usar imagens do Docker do Docker Hub e do Amazon Elastic Container Registry (Amazon ECR) para definir dependências de ambiente e biblioteca.

  Configure o Docker e execute aplicações do Spark com o Docker usando o Amazon EMR 6.x.
- O EMR é compatível com as transações ACID do Apache Hive: o Amazon EMR 6.1.0 adiciona suporte às transações ACID do Hive para estar em conformidade com as propriedades ACID de um banco de dados. Com esse atributo, você pode executar as operações `INSERT, UPDATE, DELETE,` e `MERGE` em tabelas gerenciadas do Hive com dados no Amazon Simple Storage Service (Amazon S3). Esse é um atributo essencial para casos de uso, como ingestão de streaming, redefinição de dados, atualizações em massa usando MERGE e mudanças lentas de dimensões. Para obter mais informações, incluindo exemplos de configuração e casos de uso, consulte O Amazon EMR é compatível com as transações ACID do Apache Hive.

Alterações, melhorias e problemas resolvidos

- Esta é uma versão para corrigir problemas com o Amazon EMR Scaling quando ele não consegue up/scale reduzir a escala de um cluster com sucesso ou causa falhas no aplicativo.

- Corrigido um problema em que as solicitações de escalabilidade falhavam em um cluster grande e altamente utilizado quando os daemons do Amazon EMR no cluster estavam executando atividades de verificação de integridade, como a coleta do estado do nó do YARN e o estado do nó do HDFS. Isso estava acontecendo porque os daemons no cluster não conseguiam comunicar os dados do status de integridade de um nó aos componentes internos do Amazon EMR.
- Aprimorados os daemons do EMR no cluster para rastrear corretamente os estados dos nós quando são reutilizados endereços IP para melhorar a confiabilidade durante operações de escalabilidade.
- SPARK-29683. Corrigido um problema em que ocorriam falhas de trabalho durante a redução da escala verticalmente do cluster, pois o Spark presumia que todos os nós disponíveis estavam na lista de negação.
- YARN-9011. Corrigido um problema em que ocorriam falhas de trabalho devido a uma condição de corrida na desativação do YARN quando o cluster tentava aumentar ou reduzir a escala verticalmente.
- Corrigido problema com falhas de etapas ou tarefas durante a escalabilidade do cluster ao ser garantido que os estados dos nós fossem sempre consistentes entre os daemons do Amazon EMR no cluster e o YARN/HDFS.
- Corrigido um problema em que operações de cluster, como redução de escala verticalmente e envio de etapas, falhavam para clusters do Amazon EMR habilitados com a autenticação Kerberos. Isso ocorreu porque o daemon no cluster do Amazon EMR não renovou o tíquete Kerberos, que é necessário para se comunicar com segurança com a execução no nó primário. HDFS/YARN
- As versões mais recentes do Amazon EMR corrigem o problema com um limite menor de “Máximo de arquivos abertos” em relação às versões mais antigas AL2 no Amazon EMR. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR agora incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”.
- O Apache Flink não é compatível com o EMR 6.0.0, mas é compatível com o EMR 6.1.0 com o Flink 1.11.0. Esta é a primeira versão do Flink a oficialmente oferecer suporte ao Hadoop 3. Consulte o Anúncio de versão do Apache Flink 1.11.0.
- O Ganglia foi removido dos pacotes padrão do EMR 6.1.0.

Problemas conhecidos

- Limite inferior de “Máximo de arquivos abertos” em versões mais antigas AL2 [corrigido em versões mais recentes]. Versões do Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 e emr-6.2.0 são baseadas em versões mais antigas do Amazon Linux 2 (), que AL2 têm uma configuração de limite inferior para “Máximo de arquivos abertos” quando clusters do Amazon EMR são criados com a AMI padrão. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”. Versões com o limite inferior de arquivos abertos causam o erro “Muitos arquivos abertos” ao ser enviado um trabalho do Spark. Nas versões afetadas, a AMI padrão do Amazon EMR tem uma configuração de ulimit padrão de 4096 para “Máximo de arquivos abertos”, que é inferior ao limite de 65536 arquivos na AMI mais recente do Amazon Linux 2. A configuração inferior de ulimit para “Máximo de arquivos abertos” causa falhas em trabalhos do Spark quando o driver e o executor do Spark tentam abrir mais de 4096 arquivos. Para corrigir o problema, o Amazon EMR tem um script de ação de bootstrap (BA) que ajusta a configuração de ulimit na criação do cluster.

  Se você está usando uma versão mais antiga do Amazon EMR que não tem a correção permanente para esse problema, a solução alternativa a seguir permite que você defina explicitamente o ulimit instance-controller para um máximo de 65536 arquivos.

  Defina explicitamente um ulimit na linha de comando

  1. Edite `/etc/systemd/system/instance-controller.service` para adicionar os seguintes parâmetros à seção Serviço.

     ```
     LimitNOFILE=65536

     LimitNPROC=65536
     ```

  2. Reiniciar InstanceController

     ```
     $ sudo systemctl daemon-reload

     $ sudo systemctl restart instance-controller
     ```

  Defina um ulimit usando a ação de bootstrap (BA)

  Você também pode usar um script de ação de bootstrap (BA) para configurar o ulimit instance-controller para 65536 arquivos na criação do cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- **Important**

  O Amazon EMR 6.1.0 e 6.2.0 incluem um problema de performance que pode afetar criticamente todas as operações de inserção, upsert e exclusão do Hudi. Se você planeja usar o Hudi com o Amazon EMR 6.1.0 ou 6.2.0, entre em AWS contato com o suporte para obter um Hudi RPM corrigido.

- Se você definir uma configuração personalizada de coleta de lixo com `spark.driver.extraJavaOptions` `espark.executor.extraJavaOptions`, isso resultará em falha de driver/executor inicialização com o EMR 6.1 devido à configuração conflitante da coleta de lixo. Em vez disso, com o EMR versão 6.1.0, você deve especificar uma configuração personalizada de coleta de resíduos do Spark para drivers e executores com as propriedades `spark.driver.defaultJavaOptions` e `spark.executor.defaultJavaOptions`. Leia mais em Ambiente de runtime do Apache Spark e Configurar a coleta de resíduos do Spark no Amazon EMR 6.1.0.
- Usar o Pig com o Oozie (e dentro do Hue, já que o Hue usa ações do Oozie para executar scripts do Pig) gera um erro em que uma biblioteca nativa lzo não pode ser carregada. Essa mensagem de erro é informativa e não impede a execução do Pig.
- Suporte de simultaneidade do Hudi: atualmente, o Hudi não é compatível com gravações simultâneas em uma única tabela do Hudi. Além disso, o Hudi reverte todas as alterações

feitas por gravadores em andamento antes de permitir que um novo gravador seja iniciado. As gravações simultâneas podem interferir nesse mecanismo e introduzir condições de corrida, o que pode causar corrupção de dados. Você deve garantir que, como parte do seu fluxo de trabalho de processamento de dados, só exista um gravador do Hudi operando em uma tabela do Hudi em qualquer instante. O Hudi permite vários leitores simultâneos operando na mesma tabela do Hudi.

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

  Solução:

  - SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
  - Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

- Há um problema no Amazon EMR 6.1.0 que afeta os clusters que executam o Presto. Depois de um longo período (dias), o cluster pode gerar erros, como “su: failed to execute /bin/bash: Resource temporarily unavailable” ou “shell request failed on channel 0”. Esse problema é causado por um processo interno do Amazon EMR (InstanceController) que está gerando muitos processos leves (LWP), o que acaba fazendo com que o usuário do Hadoop exceda seu limite de nproc.

Isso impede que o usuário abra processos adicionais. A solução para esse problema é fazer a atualização para o EMR 6.2.0.

Versão 6.0.0

As notas da versão a seguir incluem informações para a versão 6.0.0 do Amazon EMR.

Data da versão inicial: 10 de março de 2020

Aplicações compatíveis

- AWS SDK para Java versão 1.11.711
- Ganglia versão 3.7.2
- Hadoop versão 3.2.1
- HBase versão 2.2.3
- HCatalog versão 3.1.2
- Hive versão 3.1.2
- Hudi versão 0.5.0 incubadora
- Hue versão 4.4.0
- JupyterHub versão 1.0.0
- Livy versão 0.6.0
- MXNet versão 1.5.1
- Oozie versão 5.1.0
- Phoenix versão 5.0.0
- Presto versão 0.230
- Spark versão 2.4.4
- TensorFlow versão 1.14.0
- Zeppelin versão 0.9.0-SNAPSHOT
- Zookeeper versão 3.4.14
- Conectores e drivers: DynamoDB Connector 4.14.0

> **Note**
>
> Flink, Sqoop, Pig e Mahout não estão disponíveis na versão 6.0.0 do Amazon EMR.

Novos recursos

- Suporte ao runtime do Docker do YARN: aplicações do YARN, como trabalhos do Spark, agora podem ser executados no contexto de um contêiner do Docker. Isso permite que você defina facilmente dependências em uma imagem do Docker sem a necessidade de instalar bibliotecas personalizadas no cluster do Amazon EMR. Para obter mais informações, consulte Configurar a integração do Docker e Executar aplicações do Spark com o Docker usando o Amazon EMR 6.0.0.
- Suporte ao LLAP do Hive – agora o Hive oferece suporte ao modo de execução do LLAP para melhorar o desempenho da consulta. Para obter mais informações, consulte Usar o LLAP do Hive.

Alterações, melhorias e problemas resolvidos

- Esta é uma versão para corrigir problemas com o Amazon EMR Scaling quando ele não consegue up/scale reduzir a escala de um cluster com sucesso ou causa falhas no aplicativo.
- Corrigido um problema em que as solicitações de escalabilidade falhavam em um cluster grande e altamente utilizado quando os daemons do Amazon EMR no cluster estavam executando atividades de verificação de integridade, como a coleta do estado do nó do YARN e o estado do nó do HDFS. Isso estava acontecendo porque os daemons no cluster não conseguiam comunicar os dados do status de integridade de um nó aos componentes internos do Amazon EMR.
- Aprimorados os daemons do EMR no cluster para rastrear corretamente os estados dos nós quando são reutilizados endereços IP para melhorar a confiabilidade durante operações de escalabilidade.
- SPARK-29683. Corrigido um problema em que ocorriam falhas de trabalho durante a redução da escala verticalmente do cluster, pois o Spark presumia que todos os nós disponíveis estavam na lista de negação.
- YARN-9011. Corrigido um problema em que ocorriam falhas de trabalho devido a uma condição de corrida na desativação do YARN quando o cluster tentava aumentar ou reduzir a escala verticalmente.
- Corrigido problema com falhas de etapas ou tarefas durante a escalabilidade do cluster ao ser garantido que os estados dos nós fossem sempre consistentes entre os daemons do Amazon EMR no cluster e o YARN/HDFS.

- Corrigido um problema em que operações de cluster, como redução de escala verticalmente e envio de etapas, falhavam para clusters do Amazon EMR habilitados com a autenticação Kerberos. Isso ocorreu porque o daemon no cluster do Amazon EMR não renovou o tíquete Kerberos, que é necessário para se comunicar com segurança com a execução no nó primário. HDFS/YARN
- As versões mais recentes do Amazon EMR corrigem o problema com um limite menor de “Máximo de arquivos abertos” em relação às versões mais antigas AL2 no Amazon EMR. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR agora incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”.
- Amazon Linux
  - O Amazon Linux 2 é o sistema operacional da série 6.x do EMR.
  - `systemd` é usado para gerenciamento de serviço em vez de `upstart`, usado no Amazon Linux 1.
- Java Development Kit (JDK)
  - O JDK 8 do Coretto é o JDK padrão da série 6.x do EMR.
- Scala
  - O Scala 2.12 é usado com o Apache Spark e com o Apache Livy.
- Python 3
  - Agora o Python 3 é a versão padrão do Python no EMR.
- Rótulos de nó do YARN
  - A partir do Amazon EMR série 6.x, o recurso de rótulos de nó do YARN é desabilitado por padrão. Os principais processos do aplicativo podem ser executados tanto nos nós core como nos nós de tarefa por padrão. É possível habilitar o recurso de rótulos de nó do YARN configurando as seguintes propriedades: `yarn.node-labels.enabled` e `yarn.node-labels.am.default-node-label-expression`. Para obter mais informações, consulte Noções básicas sobre nós de tarefa, centrais e primários.

Problemas conhecidos

- Limite inferior de “Máximo de arquivos abertos” em versões mais antigas AL2 [corrigido em versões mais recentes]. Versões do Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 e emr-6.2.0 são baseadas em versões mais antigas do Amazon Linux 2 (), que AL2 têm uma configuração de limite inferior para “Máximo de arquivos abertos” quando clusters do

Amazon EMR são criados com a AMI padrão. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”. Versões com o limite inferior de arquivos abertos causam o erro “Muitos arquivos abertos” ao ser enviado um trabalho do Spark. Nas versões afetadas, a AMI padrão do Amazon EMR tem uma configuração de ulimit padrão de 4096 para “Máximo de arquivos abertos”, que é inferior ao limite de 65536 arquivos na AMI mais recente do Amazon Linux 2. A configuração inferior de ulimit para “Máximo de arquivos abertos” causa falhas em trabalhos do Spark quando o driver e o executor do Spark tentam abrir mais de 4096 arquivos. Para corrigir o problema, o Amazon EMR tem um script de ação de bootstrap (BA) que ajusta a configuração de ulimit na criação do cluster.

Se você está usando uma versão mais antiga do Amazon EMR que não tem a correção permanente para esse problema, a solução alternativa a seguir permite que você defina explicitamente o ulimit instance-controller para um máximo de 65536 arquivos.

Defina explicitamente um ulimit na linha de comando

1. Edite `/etc/systemd/system/instance-controller.service` para adicionar os seguintes parâmetros à seção Serviço.

   ```
   LimitNOFILE=65536

   LimitNPROC=65536
   ```

2. Reiniciar InstanceController

   ```
   $ sudo systemctl daemon-reload

   $ sudo systemctl restart instance-controller
   ```

Defina um ulimit usando a ação de bootstrap (BA)

Você também pode usar um script de ação de bootstrap (BA) para configurar o ulimit instance-controller para 65536 arquivos na criação do cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
```

```
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- O shell interativo do Spark PySpark, incluindo o SparkR e o spark-shell, não é compatível com o uso do Docker com bibliotecas adicionais.
- Para usar o Python 3 com a versão 6.0.0 do Amazon EMR, adicione `PATH` a `yarn.nodemanager.env-whitelist`.
- A funcionalidade Live Long and Process (LLAP) não é suportada quando você usa o AWS Glue Data Catalog como metastore do Hive.
- Ao usar o Amazon EMR 6.0.0 com a integração do Spark e do Docker, você precisa configurar as instâncias no cluster com o mesmo tipo de instância e a mesma quantidade de volumes do EBS para evitar falhas quando enviar um trabalho do Spark com runtime do Docker.
- No Amazon EMR 6.0.0, no Amazon HBase S3, o modo de armazenamento é afetado pelo problema do HBASE-24286. HBase o master não pode inicializar quando o cluster é criado usando dados existentes do S3.
- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

  Solução:

  - SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.

- Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do hadoop.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, o arquivo keytab está localizado em /etc/hadoop.keytab e a entidade principal está na forma de hadoop/<hostname>@<REALM>.

> **Note**
>
> Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versão 5.30.1

As notas da versão a seguir incluem informações para a versão 5.30.1 do Amazon EMR. As alterações são referentes à versão 5.30.0.

Data da versão inicial: 30 de junho de 2020

Data da última atualização: 24 de agosto de 2020

Alterações, melhorias e problemas resolvidos

- As versões mais recentes do Amazon EMR corrigem o problema com um limite menor de “Máximo de arquivos abertos” em relação às versões mais antigas AL2 no Amazon EMR. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR agora incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”.
- Corrigido um problema em que o processo do controlador da instância gerava um número infinito de processos.
- Corrigido um problema em que o Hue não conseguia executar uma consulta do Hive, mostrando a mensagem “o banco de dados está bloqueado” e impedindo a execução de consultas.
- Corrigido um problema no Spark para permitir que mais tarefas fossem executadas simultaneamente no cluster do EMR.

- Corrigido um problema no caderno Jupyter que causava um “erro de muitos arquivos abertos” no servidor Jupyter.
- Corrigido um problema com as horas de início do cluster.

Novos recursos

- A interface do usuário do Tez e as interfaces de aplicações persistentes do servidor de linha do tempo do YARN estão disponíveis com as versões 6.x do Amazon EMR e com as versões 5.30.1 e posteriores do EMR. O acesso com um clique no link do histórico persistente da aplicação permite que você acesse rapidamente o histórico de tarefas sem configurar um proxy da Web por meio de uma conexão SSH. Os logs de clusters ativos e encerrados ficam disponíveis por 30 dias após o término da aplicação. Para obter mais informações, consulte View Persistent Application User Interfaces no Guia de gerenciamento do Amazon EMR.
- A execução do EMR Notebook APIs está disponível para executar notebooks EMR por meio de um script ou linha de comando. A capacidade de iniciar, parar, listar e descrever as execuções do notebook EMR sem o AWS console permite que você controle programaticamente um notebook EMR. Ao usar uma célula do caderno parametrizada, você pode passar valores de parâmetros diferentes para um caderno sem precisar criar uma cópia do caderno para cada novo conjunto de valores de parâmetros. Consulte Ações de API do EMR. Para obter um exemplo de código, consulte Exemplos de comandos para executar Cadernos do EMR programaticamente.

Problemas conhecidos

- Limite inferior de “Máximo de arquivos abertos” em versões mais antigas AL2 [corrigido em versões mais recentes]. Versões do Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 e emr-6.2.0 são baseadas em versões mais antigas do Amazon Linux 2 (), que AL2 têm uma configuração de limite inferior para “Máximo de arquivos abertos” quando clusters do Amazon EMR são criados com a AMI padrão. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”. Versões com o limite inferior de arquivos abertos causam o erro “Muitos arquivos abertos” ao ser enviado um trabalho do Spark. Nas versões afetadas, a AMI padrão do Amazon EMR tem uma configuração de ulimit padrão de 4096 para “Máximo de arquivos abertos”, que é inferior ao limite de 65536 arquivos na AMI mais recente do Amazon Linux 2. A configuração inferior de ulimit para “Máximo de arquivos abertos” causa falhas em trabalhos do Spark quando o driver e o executor do Spark tentam abrir mais de

4096 arquivos. Para corrigir o problema, o Amazon EMR tem um script de ação de bootstrap (BA) que ajusta a configuração de ulimit na criação do cluster.

Se você está usando uma versão mais antiga do Amazon EMR que não tem a correção permanente para esse problema, a solução alternativa a seguir permite que você defina explicitamente o ulimit instance-controller para um máximo de 65536 arquivos.

Defina explicitamente um ulimit na linha de comando

1. Edite `/etc/systemd/system/instance-controller.service` para adicionar os seguintes parâmetros à seção Serviço.

   ```
   LimitNOFILE=65536

   LimitNPROC=65536
   ```

2. Reiniciar InstanceController

   ```
   $ sudo systemctl daemon-reload

   $ sudo systemctl restart instance-controller
   ```

Defina um ulimit usando a ação de bootstrap (BA)

Você também pode usar um script de ação de bootstrap (BA) para configurar o ulimit instance-controller para 65536 arquivos na criação do cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
```

```
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- Cadernos do EMR

  O atributo que permite instalar kernels e bibliotecas Python adicionais no nó primário do cluster está desabilitado por padrão na versão 5.30.1 do EMR. Para obter mais informações sobre esse atributo, consulte Instalar kernels e bibliotecas Python em um nó primário do cluster.

  Para habilitar o recurso, faça o seguinte:

  1. Certifique-se de que a política de permissões anexada ao perfil de serviço para os Cadernos do EMR permite a seguinte ação:

     ```
     elasticmapreduce:ListSteps
     ```

     Para obter mais informações, consulte Função de serviço do EMR Notebooks.

  2. Use o AWS CLI para executar uma etapa no cluster que configura os Notebooks EMR, conforme mostrado no exemplo a seguir. *us-east-1*Substitua pela região em que seu cluster reside. Para obter mais informações, consulte Adding Steps to a Cluster Using the AWS CLI.

     ```
     aws emr add-steps  --cluster-id MyClusterID --steps
      Type=CUSTOM_JAR,Name=EMRNotebooksSetup,ActionOnFailure=CONTINUE,Jar=s3://us-
     east-1.elasticmapreduce/libs/script-runner/script-runner.jar,Args=["s3://
     awssupportdatasvcs.com/bootstrap-actions/EMRNotebooksSetup/emr-notebooks-setup.sh"]
     ```

- Ajuste de escala gerenciado

  As operações de ajuste de escala gerenciado nos clusters das versões 5.30.0 e 5.30.1 sem o Presto instalado podem causar falhas na aplicação ou fazer com que um grupo de instâncias ou uma frota de instâncias uniforme permaneça no estado ARRESTED, sobretudo quando uma operação de redução da escala verticalmente logo é seguida por uma operação de aumento da escala verticalmente.

  Como solução alternativa, escolha o Presto como uma aplicação a ser instalada ao criar um cluster com as versões 5.30.0 e 5.30.1 do Amazon EMR, mesmo que o trabalho não exija o Presto.

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

Solução:

- SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
- Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
  - O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que

existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.

Versão 5.30.0

As notas da versão a seguir incluem informações para a versão 5.30.0 do Amazon EMR. As alterações são referentes à versão 5.29.0.

Data da versão inicial: 13 de maio de 2020

Data da última atualização: 25 de junho de 2020

Atualizações

- Atualizado para AWS SDK para Java a versão 1.11.759
- SDK do Amazon SageMaker Spark atualizado para a versão 1.3.0
- Atualização do EMR Record Server para a versão 1.6.0
- Atualização do Flink para a versão 1.10.0
- Atualização do Ganglia para a versão 3.7.2
- Atualizado HBase para a versão 1.4.13
- Atualização do Hudi para a versão 0.5.2-incubating
- Atualização do Hue para a versão 4.6.0
- Atualizado JupyterHub para a versão 1.1.0
- Atualização do Livy para a versão 0.7.0-incubating
- Atualização do Oozie para a versão 5.2.0
- Atualização do Presto para a versão 0.232
- Atualização do Spark para a versão 2.4.5
- Atualização de conectores e drivers: Amazon Glue Connector 1.12.0; Amazon Kinesis Connector 3.5.0; EMR DynamoDB Connector 4.14.0

Novos recursos

- Cadernos do EMR: quando usados com clusters do EMR criados com a versão 5.30.0, os kernels dos Cadernos do EMR são executados no cluster. Isso melhora o desempenho do bloco de anotações e permite que instalar e personalizar kernels. Você também pode instalar bibliotecas Python no nó primário do cluster. Para obter mais informações, consulte Instalar e usar kernels e bibliotecas no EMR Management Guide.
- Ajuste de escala gerenciado: com as versões 5.30.0 e posteriores do Amazon EMR, é possível habilitar o ajuste de escala gerenciado pelo EMR para aumentar ou diminuir automaticamente o número de instâncias ou unidades no cluster com base na workload. O Amazon EMR avalia continuamente as métricas do cluster para tomar decisões de ajuste de escala que otimizam os clusters em termos de custo e velocidade. Para obter mais informações, consulte Escalar recursos de cluster no Guia de gerenciamento do Amazon EMR.
- Criptografe arquivos de log armazenados no Amazon S3 — Com o Amazon EMR versão 5.30.0 e posterior, você pode criptografar arquivos de log armazenados no Amazon S3 com uma chave gerenciada pelo cliente. AWS KMS Para obter mais informações, consulte Encrypt log files stored in Amazon S3 no Guia de gerenciamento do Amazon EMR.
- Suporte ao Amazon Linux 2: nas versões 5.30.0 e posteriores do EMR, o EMR usa o sistema operacional Amazon Linux 2. A nova personalização AMIs (Amazon Machine Image) deve ser baseada na Amazon Linux 2 AMI. Para obter mais informações, consulte Usando uma AMI personalizada.
- Escalabilidade automática tranquila do Presto: os clusters do EMR que usam a versão 5.30.0 podem ser definidos com um período limite de ajuste de escala automático que dá tempo às tarefas do Presto para concluir a execução antes do nó delas ser desativado. Para obter mais informações, consulte Usar a escalabilidade automática do Presto com desativação tranquila.
- Criação de instância de frota com nova opção de estratégia de alocação: uma nova opção de estratégia de alocação está disponível nas versões 5.12.1 e posteriores do EMR. Ele oferece provisionamento de cluster mais rápido, alocação de spot mais precisa e menos interrupção de instâncias spot. São necessárias atualizações para perfis de serviço do EMR não padrão. Consulte Configurar frotas de instâncias.
- Comandos sudo systemctl stop e sudo systemctl start: nas versões 5.30.0 e posteriores do EMR, que usam o sistema operacional Amazon Linux 2, o EMR usa os comandos `sudo systemctl stop` e `sudo systemctl start` para reiniciar serviços. Para obter mais informações, consulte Como reinicio um serviço no Amazon EMR?.

Alterações, melhorias e problemas resolvidos

- O EMR versão 5.30.0 não instala o Ganglia por padrão. É possível selecionar explicitamente o Ganglia para ser instalado ao criar um cluster.
- Otimizações do desempenho do Spark
- Otimizações do desempenho do Presto
- O Python 3 é o padrão para as versões 5.30.0 e posteriores do Amazon EMR.
- O grupo de segurança gerenciado padrão para acesso ao serviço em sub-redes privadas foi atualizado com novas regras. Se você usar um grupo de segurança personalizado para acesso ao serviço, será necessário incluir as mesmas regras do grupo de segurança gerenciado padrão. Para obter mais informações, consulte Grupo de segurança gerenciado pelo Amazon EMR para acesso ao serviço (sub-redes privadas). Se você usar um perfil de serviço personalizado para o Amazon EMR, será necessário conceder permissão para `ec2:describeSecurityGroups` de modo que o EMR possa validar se os grupos de segurança são criados corretamente. Se você usar o `EMR_DefaultRole`, essa permissão já estará incluída na política gerenciada padrão.

Problemas conhecidos

- Limite inferior de “Máximo de arquivos abertos” em versões mais antigas AL2 [corrigido em versões mais recentes]. Versões do Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 e emr-6.2.0 são baseadas em versões mais antigas do Amazon Linux 2 (), que AL2 têm uma configuração de limite inferior para “Máximo de arquivos abertos” quando clusters do Amazon EMR são criados com a AMI padrão. As versões 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 e posteriores do Amazon EMR incluem uma correção permanente com uma configuração mais alta de “Máximo de arquivos abertos”. Versões com o limite inferior de arquivos abertos causam o erro “Muitos arquivos abertos” ao ser enviado um trabalho do Spark. Nas versões afetadas, a AMI padrão do Amazon EMR tem uma configuração de ulimit padrão de 4096 para “Máximo de arquivos abertos”, que é inferior ao limite de 65536 arquivos na AMI mais recente do Amazon Linux 2. A configuração inferior de ulimit para “Máximo de arquivos abertos” causa falhas em trabalhos do Spark quando o driver e o executor do Spark tentam abrir mais de 4096 arquivos. Para corrigir o problema, o Amazon EMR tem um script de ação de bootstrap (BA) que ajusta a configuração de ulimit na criação do cluster.

  Se você está usando uma versão mais antiga do Amazon EMR que não tem a correção permanente para esse problema, a solução alternativa a seguir permite que você defina explicitamente o ulimit instance-controller para um máximo de 65536 arquivos.

Defina explicitamente um ulimit na linha de comando

1. Edite `/etc/systemd/system/instance-controller.service` para adicionar os seguintes parâmetros à seção Serviço.

   ```
   LimitNOFILE=65536
   ```

   ```
   LimitNPROC=65536
   ```

2. Reiniciar InstanceController

   ```
   $ sudo systemctl daemon-reload
   ```

   ```
   $ sudo systemctl restart instance-controller
   ```

Defina um ulimit usando a ação de bootstrap (BA)

Você também pode usar um script de ação de bootstrap (BA) para configurar o ulimit instance-controller para 65536 arquivos na criação do cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- Ajuste de escala gerenciado

As operações de ajuste de escala gerenciado nos clusters das versões 5.30.0 e 5.30.1 sem o Presto instalado podem causar falhas na aplicação ou fazer com que um grupo de instâncias ou uma frota de instâncias uniforme permaneça no estado `ARRESTED`, sobretudo quando uma operação de redução da escala verticalmente logo é seguida por uma operação de aumento da escala verticalmente.

Como solução alternativa, escolha o Presto como uma aplicação a ser instalada ao criar um cluster com as versões 5.30.0 e 5.30.1 do Amazon EMR, mesmo que o trabalho não exija o Presto.

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

  Solução:

  - SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
  - Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

- O mecanismo de banco de dados padrão para o Hue 4.6.0 é SQLite, o que causa problemas quando você tenta usar o Hue com um banco de dados externo. Para corrigir isso, defina `engine`

na sua classificação de configuração `hue-ini` como `mysql`. Esse problema foi corrigido na versão 5.30.1 do Amazon EMR.

- Quando você usa o Spark com a formatação de localização de partições do Hive para ler dados no Amazon S3 e executa o Spark nas versões 5.30.0 a 5.36.0 e 6.2.0 a 6.9.0 do Amazon EMR, pode encontrar um problema que impede que o cluster leia os dados corretamente. Isso poderá acontecer se suas partições tiverem todas as características a seguir:
  - Duas ou mais partições são verificadas na mesma tabela.
  - Pelo menos um caminho de diretório de partição é um prefixo de pelo menos outro caminho de diretório de partição, por exemplo, `s3://bucket/table/p=a` é um prefixo de `s3://bucket/table/p=a b`.
  - O primeiro caractere que segue o prefixo no outro diretório de partição tem um valor UTF-8 menor que o caractere `/` (U+002F). Por exemplo, o caractere de espaço (U+0020) que ocorre entre a e b em `s3://bucket/table/p=a b` se enquadra nessa categoria. Observe que existem 14 outros caracteres que não são de controle: `!"#$%&'()*+,-`. Para obter mais informações, consulte Tabela de codificação UTF-8 e caracteres Unicode.

  Como solução alternativa para esse problema, defina a configuração `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` na classificação `spark-defaults`.

Versão 5.29.0

As notas da versão a seguir incluem informações para a versão 5.29.0 do Amazon EMR. As alterações são referentes à versão 5.28.1.

Data da versão inicial: 17 de janeiro de 2020

Atualizações

- Atualizado para AWS SDK para Java a versão 1.11.682
- Atualizado o Hive para a versão 2.3.6
- Atualizado o Flink para a versão 1.9.1
- Atualizado o EMRFS para a versão 2.38.0
- Atualizado o conector do DynamoDB para EMR, versão 4.13.0

Alterações, melhorias e problemas resolvidos

- Spark
  - Otimizações do desempenho do Spark
- EMRFS
  - O Guia de gerenciamento é atualizado para as configurações padrão emrfs-site.xml para uma visualização consistente.

Problemas conhecidos

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

  Solução:

  - SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
  - Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versão 5.28.1

As notas da versão a seguir incluem informações para a versão 5.28.1 do Amazon EMR. As alterações são referentes à versão 5.28.0.

Data da versão inicial: 10 de janeiro de 2020

Alterações, melhorias e problemas resolvidos

- Spark
  - Correção de problemas de compatibilidade do Spark.
- CloudWatch Métricas
  - Foi corrigida a publicação do Amazon CloudWatch Metrics em um cluster do EMR com vários nós primários.
- Desabilitada mensagem de log
  - Desabilitada mensagem de log falsa, “... uso de versão antiga (<4.5.8) do cliente Apache http”.

Problemas conhecidos

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

  Solução:

  - SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
  - Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versão 5.28.0

As notas da versão a seguir incluem informações para a versão 5.28.0 do Amazon EMR. As alterações são referentes à 5.27.0.

Data da versão inicial: 12 de novembro de 2019

Atualizações

- Flink atualizado para a versão 1.9.0
- Atualizado o Hive para a versão 2.3.6
- Atualizado MXNet para a versão 1.5.1
- Phoenix atualizado para a versão 4.14.3
- Presto atualizado para a versão 0.227
- Zeppelin atualizado para a versão 0.8.2

Novos recursos

- O Apache Hudi agora está disponível para o Amazon EMR instalar na criação de um cluster. Para obter mais informações, consulte Hudi.
- (25 de novembro de 2019) Agora você pode optar por executar várias etapas em paralelo para melhorar a utilização do cluster e economizar custos. Pode também cancelar etapas pendentes e em execução. Para obter mais informações, consulte Trabalhar com etapas usando o console AWS CLI e.

- (3 de dezembro de 2019) Agora você pode criar e executar clusters do EMR no. AWS Outposts AWS Outposts habilita AWS serviços, infraestrutura e modelos operacionais nativos em instalações locais. Em AWS Outposts ambientes, você pode usar as mesmas AWS APIs ferramentas e infraestrutura que usa na AWS nuvem. Para obter mais informações, consulte Clusters do EMR ativados. AWS Outposts
- (11 de março de 2020) A partir da versão 5.28.0 do Amazon EMR, você pode criar e executar clusters do Amazon EMR em uma sub-rede de Zonas AWS Locais como uma extensão lógica de uma região que suporta Zonas Locais. AWS Uma zona local permite que os recursos do Amazon EMR e um subconjunto de AWS serviços, como serviços de computação e armazenamento, estejam localizados mais perto dos usuários, fornecendo acesso de latência muito baixa aos aplicativos executados localmente. Para obter uma lista das zonas locais disponíveis, consulte Zonas locais da AWS. Para obter informações sobre como acessar as Zonas AWS Locais disponíveis, consulte Regiões, Zonas de Disponibilidade e Zonas Locais.

  Atualmente, as zonas locais não são compatíveis com os Cadernos do Amazon EMR nem com conexões diretas com o Amazon EMR usando o endpoint da VPC da interface (AWS PrivateLink).

Alterações, melhorias e problemas resolvidos

- Suporte expandido do aplicativo para clusters de alta disponibilidade
  - Para obter mais informações, consulte Supported applications in an EMR cluster with Multiple Primary Nodes no Guia de gerenciamento do Amazon EMR.
- Spark
  - Otimizações da performance
- Hive
  - Otimizações da performance
- Presto
  - Otimizações da performance

Problemas conhecidos

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução

por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

Solução:

- SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
- Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versão 5.27.0

As notas da versão a seguir incluem informações para a versão 5.27.0 do Amazon EMR. As alterações são referentes à versão 5.26.0.

Data da versão inicial: 23 de setembro de 2019

Atualizações

- AWS SDK para Java 1.11.615
- Flink 1.8.1
- JupyterHub 1.0.0
- Spark 2.4.4
- Tensorflow 1.14.0

- Conectores e drivers:
  - Conector do DynamoDB 4.12.0

Novos recursos

- (24 de outubro de 2019) Os seguintes novos atributos dos Cadernos do EMR estão disponíveis em todas as versões do Amazon EMR.
  - É possível associar repositórios do Git a Cadernos do EMR para armazenar os cadernos em um ambiente controlado de versão. Você pode compartilhar códigos com pares e reutilizar cadernos Jupyter existentes por meio de repositórios do Git remotos. Para obter mais informações, consulte Associate Git Repositories with Amazon EMR Notebooks no Guia de gerenciamento do Amazon EMR.
  - O utilitário nbdime agora está disponível em Cadernos do EMR para simplificar a comparação e a mesclagem de cadernos.
  - Agora há suporte para notebooks EMR. JupyterLab JupyterLab é um ambiente de desenvolvimento interativo baseado na Web totalmente compatível com os notebooks Jupyter. Agora você pode optar por abrir seu caderno em um JupyterLab ou no editor de cadernos Jupyter.
- (30 de outubro de 2019) Com as versões 5.25.0 e posteriores do Amazon EMR, é possível conectar-se ao servidor de histórico do Spark na página Resumo do cluster ou na guia Histórico da aplicação no console. Em vez de configurar um proxy da Web por meio de uma conexão SSH, você pode acessar rapidamente a interface do usuário do servidor de histórico do Spark para visualizar as métricas da aplicação e acessar arquivos de log relevantes para clusters ativos e encerrados. Para obter mais informações, consulte Acesso fora do cluster a interfaces de usuário de aplicações persistentes no Guia de gerenciamento do Amazon EMR.

Alterações, melhorias e problemas resolvidos

- Cluster do Amazon EMR com vários nós primários
  - Você pode instalar e executar o Flink em um cluster do Amazon EMR com vários nós primários. Para obter mais informações, consulte Supported applications and features.
  - Você pode configurar a criptografia transparente do HDFS em um cluster do Amazon EMR com vários nós primários. Para obter mais informações, consulte Criptografia transparente do HDFS em clusters do EMR com vários nós primários.

  - Agora você pode modificar a configuração das aplicações em execução em um cluster do Amazon EMR com vários nós primários. Para obter mais informações, consulte Supplying a Configuration for an Instance Group in a Running Cluster.
- Conector Amazon EMR-DynamoDB
  - O conector Amazon EMR-DynamoDB agora é compatível com os seguintes tipos de dados do DynamoDB: booleanos, lista, mapa, item, nulos. Para obter mais informações, consulte Configurar uma tabela do Hive para executar comandos do Hive.

Problemas conhecidos

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

  Solução:

  - SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
  - Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versão 5.26.0

As notas da versão a seguir incluem informações para a versão 5.26.0 do Amazon EMR. As alterações são referentes à versão 5.25.0.

Data da versão inicial: 8 de agosto de 2019

Data da última atualização: 19 de agosto de 2019

Atualizações

- AWS SDK para Java 1.11.595
- HBase 1.4.10
- Phoenix 4.14.2
- Conectores e drivers:
  - Conector do DynamoDB 4.11.0
  - Conector do MariaDB 2.4.2
  - Driver JDBC do Amazon Redshift, 1.2.32.1056

Novos recursos

- (Beta) Com a versão 5.26.0 do Amazon EMR, você pode iniciar um cluster que se integre ao Lake Formation. Essa integração fornece acesso refinado em nível de coluna a bancos de dados e tabelas no Glue Data Catalog. AWS Ela também permite logon único federado para Cadernos do EMR ou para o Apache Zeppelin em um sistema de identidade empresarial. Para obter mais informações, consulte Integração do Amazon EMR AWS Lake Formation com (Beta).
- (19 de agosto de 2019) O bloqueio de acesso público do Amazon EMR agora está disponível em todas as versões do Amazon EMR compatíveis com grupos de segurança. Bloquear o acesso público é uma configuração de toda a conta aplicada a cada AWS região. Bloquear o acesso público impede que um cluster seja iniciado quando qualquer grupo de segurança associado ao cluster tem uma regra que permite tráfego de entrada de IPv4 0.0.0.0/0 ou IPv6 : :/0 (acesso público) em uma porta, a menos que uma porta seja especificada como uma exceção. A porta 22 é uma exceção por padrão. Para obter mais informações, consulte Usar o bloqueio de acesso público do Amazon EMR no Guia de gerenciamento do Amazon EMR.

Alterações, melhorias e problemas resolvidos

- Cadernos do EMR
  - Com as versões 5.26.0 e posteriores do EMR, os Cadernos do EMR são compatíveis com as bibliotecas Python com escopo de caderno, além das bibliotecas Python padrão. Você pode instalar bibliotecas com escopo de caderno de dentro do editor de caderno sem precisar recriar um cluster ou reanexar um caderno a um cluster. As bibliotecas com escopo de caderno são criadas em um ambiente Python virtual para serem aplicadas somente à sessão de caderno atual. Isso permite isolar as dependências do caderno. Para obter mais informações, consulte Usar bibliotecas com escopo de caderno no Guia de gerenciamento do Amazon EMR.
- EMRFS
  - Você pode ativar um recurso ETag de verificação (Beta) configurando `fs.s3.consistent.metadata.etag.verification.enabled` como `true`. Com esse recurso, o EMRFS usa o Amazon ETags S3 para verificar se os objetos que estão sendo lidos são a versão mais recente disponível. Esse recurso é útil para casos de read-after-update uso em que os arquivos no Amazon S3 são sobrescritos, mantendo o mesmo nome. Atualmente, esse recurso de ETag verificação não funciona com o S3 Select. Para obter mais informações, consulte Configurar visualização consistente.
- Spark
  - As seguintes otimizações agora estão habilitadas por padrão: remoção dinâmica de partições, DISTINCT antes de INTERSECT, melhorias na inferência de estatísticas do plano SQL para JOIN seguida por consultas DISTINCT, nivelamento de subconsultas escalares, reordenamento otimizado de junções e junção com filtro de Bloom. Para obter mais informações, consulte Otimizar a performance do Spark.
  - Aprimorada a geração de código de estágio completo para Sort-Merge Join.
  - Aprimorado o fragmento de consulta e a reutilização de subconsultas.
  - Melhorias na pré-alocação de executores na inicialização do Spark.
  - As junções com filtro de Bloom não são mais aplicadas quando o lado menor da junção inclui uma dica de transmissão.
- Tez
  - Resolvido um problema com o Tez. A IU do Tez agora funciona em um cluster do Amazon EMR com vários nós primários.

Problemas conhecidos

- Os recursos aprimorados de geração de código em todo o estágio de Sort Merge Join podem aumentar a pressão de memória quando habilitados. Essa otimização melhora a performance, mas pode resultar em novas tentativas ou falhas de trabalho se `spark.yarn.executor.memoryOverheadFactor` não for ajustado para fornecer memória suficiente. Para desabilitar esse atributo, defina `spark.sql.sortMergeJoinExec.extendedCodegen.enabled` como falso.
- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

  Solução:

  - SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
  - Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versão 5.25.0

As notas da versão a seguir incluem informações para a versão 5.25.0 do Amazon EMR. As alterações são referentes à versão 5.24.1.

Data da versão inicial: 17 de julho de 2019

Data da última atualização: 30 de outubro de 2019

Amazon EMR 5.25.0

Atualizações

- AWS SDK para Java 1.11.566
- Hive 2.3.5
- Presto 0.220
- Spark 2.4.3
- TensorFlow 1.13.1
- Tez 0.9.2
- Zookeeper 3.4.14

Novos recursos

- (30 de outubro de 2019) Desde a versão 5.25.0 do Amazon EMR, é possível conectar-se ao servidor de histórico do Spark na página Resumo do cluster ou na guia Histórico da aplicação no console. Em vez de configurar um proxy da Web por meio de uma conexão SSH, você pode acessar rapidamente a interface do usuário do servidor de histórico do Spark para visualizar as métricas da aplicação e acessar arquivos de log relevantes para clusters ativos e encerrados. Para obter mais informações, consulte Acesso fora do cluster a interfaces de usuário de aplicações persistentes no Guia de gerenciamento do Amazon EMR.

Alterações, melhorias e problemas resolvidos

- Spark
  - Aprimorada a performance de algumas junções usando filtros de Bloom para pré-filtrar as entradas. A otimização é desabilitada por padrão e pode ser habilitada com a definição do parâmetro `spark.sql.bloomFilterJoin.enabled` de configuração do Spark como `true`.

- Aprimorada a performance do agrupamento por colunas do tipo string.
- Melhorou a memória padrão do executor Spark e a configuração dos núcleos dos tipos de instância R4 para clusters sem instalação. HBase
- Resolvido um problema anterior com o atributo de remoção dinâmica de partições, em que a tabela removida precisava estar no lado esquerdo da junção.
- Aprimorada a otimização do DISTINCT antes do INTERSECT para ser aplicada a casos adicionais envolvendo aliases.
- Aprimorada a inferência de estatísticas do plano SQL para JOIN seguida por consultas DISTINCT. Essa melhoria é desabilitada por padrão e pode ser habilitada pela definição do parâmetro `spark.sql.statsImprovements.enabled` de configuração do Spark como `true`. Essa otimização é exigida pelo atributo Distinct antes do Intersect e será habilitada automaticamente quando `spark.sql.optimizer.distinctBeforeIntersect.enabled` estiver definido como `true`.
- Otimizada a ordem de junção com base no tamanho da tabela e nos filtros. Essa otimização é desativada por padrão e pode ser ativada com a definição do parâmetro `spark.sql.optimizer.sizeBasedJoinReorder.enabled` de configuração do Spark como `true`.

Para obter mais informações, consulte Otimizar a performance do Spark.

- EMRFS
  - A configuração do EMRFS, `fs.s3.buckets.create.enabled`, agora está desabilitada por padrão. Por meio de testes, descobrimos que a desabilitação dessa configuração melhora a performance e evita a criação não intencional de buckets do S3. Se sua aplicação depende dessa funcionalidade, você pode habilitá-la definindo a propriedade `fs.s3.buckets.create.enabled` como `true` na classificação de configuração `emrfs-site`. Para obter informações, consulte Supplying a Configuration when Creating a Cluster.
- Melhorias na criptografia de disco local e na criptografia do S3 nas configurações de segurança (5 de agosto de 2019)
  - Separadas as configurações de criptografia do Amazon S3 das configurações de criptografia de disco local na configuração de segurança.
  - Adicionada uma opção para habilitar a criptografia do EBS com as versões 5.24.0 e posteriores. Selecionar essa opção criptografa o volume do dispositivo raiz, além dos volumes de armazenamento. As versões anteriores exigiam o uso de uma AMI personalizada para criptografar o volume do dispositivo raiz.

- Para obter mais informações, consulte Opções de criptografia no Guia de gerenciamento do Amazon EMR.

Problemas conhecidos

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

  Solução:

  - SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
  - Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versão 5.24.1

As notas da versão a seguir incluem informações para a versão 5.24.1 do Amazon EMR. As alterações são referentes à versão 5.24.0.

Data da versão inicial: 26 de junho de 2019

Alterações, melhorias e problemas resolvidos

- Atualizada a AMI padrão do Amazon Linux para Amazon EMR para incluir atualizações de segurança importantes do kernel Linux, incluindo o problema de negação de serviço do TCP SACK (AWS-2019-005).

Problemas conhecidos

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

  Solução:

  - SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
  - Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versão 5.24.0

As notas da versão a seguir incluem informações para a versão 5.24.0 do Amazon EMR. As alterações são referentes à versão 5.23.0.

Data da versão inicial: 11 de junho de 2019

Data da última atualização: 5 de agosto de 2019

Atualizações

- Flink 1.8.0
- Hue 4.4.0
- JupyterHub 0.9.6
- Livy 0.6.0
- MxNet 1.4.0
- Presto 0.219
- Spark 2.4.2
- AWS SDK para Java 1.11.546
- Conectores e drivers:
  - Conector do DynamoDB 4.9.0
  - Conector do MariaDB 2.4.1
  - Driver JDBC do Amazon Redshift, 1.2.27.1051

Alterações, melhorias e problemas resolvidos

- Spark
  - Adicionada otimização para remover partições dinamicamente. A otimização está desabilitada por padrão. Para habilitá-la, defina o parâmetro `spark.sql.dynamicPartitionPruning.enabled` de configuração do Spark como `true`.
  - Aprimorada a performance de consultas `INTERSECT`. Essa otimização está desabilitada por padrão. Para habilitá-la, defina o parâmetro `spark.sql.optimizer.distinctBeforeIntersect.enabled` de configuração do Spark como `true`.
  - Adicionada otimização para nivelar subconsultas escalares com agregados que usam a mesma relação. A otimização está desabilitada por padrão. Para habilitá-la, defina o parâmetro

`spark.sql.optimizer.flattenScalarSubqueriesWithAggregates.enabled` de configuração do Spark como `true`.

- Aprimorada a geração de código em todo o estágio.

Para obter mais informações, consulte Otimizar a performance do Spark.

- Melhorias na criptografia de disco local e na criptografia do S3 nas configurações de segurança (5 de agosto de 2019)
  - Separadas as configurações de criptografia do Amazon S3 das configurações de criptografia de disco local na configuração de segurança.
  - Adicionada uma opção para habilitar a criptografia do EBS. Selecionar essa opção criptografa o volume do dispositivo raiz, além dos volumes de armazenamento. As versões anteriores exigiam o uso de uma AMI personalizada para criptografar o volume do dispositivo raiz.
  - Para obter mais informações, consulte Opções de criptografia no Guia de gerenciamento do Amazon EMR.

Problemas conhecidos

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

  Solução:

  - SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
  - Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versão 5.23.0

As notas da versão a seguir incluem informações para a versão 5.23.0 do Amazon EMR. As alterações são referentes à versão 5.22.0.

Data da versão inicial: 1.º de abril de 2019

Data da última atualização: 30 de abril de 2019

Atualizações

- AWS SDK para Java 1.11.519

Novos recursos

- (30 de abril de 2019) Com o Amazon EMR 5.23.0 e versões posteriores, você pode iniciar um cluster com três nós principais para oferecer suporte à alta disponibilidade de aplicativos como YARN Resource Manager, HDFS, Spark, Hive e NameNode Ganglia. O nó primário não é mais um possível ponto de falha único com esse recurso. Se um dos nós primários apresenta falha, o Amazon EMR executa failover automaticamente para um nó primário em espera e substitui o nó primário com falha por um novo com as mesmas ações de configuração e bootstrap. Para obter mais informações, consulte Plan and Configure Primary Nodes.

Problemas conhecidos

- Interface do usuário do Tez (corrigida na versão 5.26.0 do Amazon EMR)

  A interface do usuário do Tez não funciona em um cluster do EMR com vários nós primários.
- Hue (corrigido na versão 5.24.0 do Amazon EMR)

- O Hue executado no Amazon EMR não é compatível com o Solr. Desde a versão 5.20.0 do Amazon EMR, um problema de configuração incorreta faz com que o Solr seja habilitado e uma mensagem de erro inofensiva semelhante à seguinte seja exibida:

  ```
  Solr server could not be contacted properly:
  HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
  port=1978): Max retries exceeded with url: /solr/admin/info/
  system?user.name=hue&doAs=administrator&wt=json (Caused by
  NewConnectionError(': Failed to establish a new connection: [Errno 111]
  Connection refused',))
  ```

  Para evitar que a mensagem de erro do Solr seja exibida:

  1. Conecte-se à linha de comando do nó primário usando SSH.
  2. Use um editor de texto para abrir o arquivo `hue.ini`. Por exemplo:

     ```
     sudo vim /etc/hue/conf/hue.ini
     ```

  3. Pesquise o termo `appblacklist` e modifique a linha para o seguinte:

     ```
     appblacklist = search
     ```

  4. Salve as alterações e reinicie o Hue, conforme mostrado no exemplo a seguir:

     ```
     sudo stop hue; sudo start hue
     ```

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

  Solução:

  - SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
  - Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versão 5.22.0

As notas da versão a seguir incluem informações para a versão 5.22.0 do Amazon EMR. As alterações são referentes à versão 5.21.0.

> **Important**
>
> A partir da versão 5.22.0 do Amazon EMR, o Amazon EMR AWS usa o Signature versão 4 exclusivamente para autenticar solicitações para o Amazon S3. As versões anteriores do Amazon EMR usam o AWS Signature versão 2 em alguns casos, a menos que as notas de lançamento indiquem que o Signature versão 4 é usado exclusivamente. Para obter mais informações, consulte Autenticação de solicitações (AWS assinatura versão 4) e Solicitações de autenticação (AWS assinatura versão 2) no Guia do desenvolvedor do Amazon Simple Storage Service.

Data da versão inicial: 20 de março de 2019

Atualizações

- Flink 1.7.1
- HBase 1.4.9
- Oozie 5.1.0
- Phoenix 4.14.1

- Zeppelin 0.8.1
- Conectores e drivers:
  - Conector do DynamoDB 4.8.0
  - Conector do MariaDB 2.2.6
  - Driver JDBC do Amazon Redshift, 1.2.20.1043

Novos recursos

- Modificou a configuração padrão do EBS para tipos de EC2 instância com armazenamento somente EBS. Ao ser criado um cluster usando as versões 5.22.0 e posteriores do Amazon EMR, a quantidade padrão de armazenamento do EBS aumenta de acordo com o tamanho da instância. Além disso, dividimos o armazenamento aumentado em vários volumes, proporcionando maior performance de IOPS. Se você quiser usar uma instância de configuração de armazenamento EBS diferente, você poderá especificá-la ao criar um cluster do EMR ou adicionar nós a um cluster existente. Para obter mais informações sobre a quantidade de armazenamento e o número de volumes alocados por padrão para cada tipo de instância, consulte Armazenamento padrão do EBS para instâncias no Guia de gerenciamento do Amazon EMR.

Alterações, melhorias e problemas resolvidos

- Spark
  - Introduzida uma nova propriedade de configuração para o Spark no YARN, `spark.yarn.executor.memory0verheadFactor`. O valor dessa propriedade é um fator de escala que define o valor da sobrecarga de memória como uma porcentagem da memória do executor, com um mínimo de 384 MB. Se a sobrecarga de memória for definida explicitamente usando `spark.yarn.executor.memory0verhead`, essa propriedade não terá efeito. O valor padrão é `0.1875`, representando 18,75%. Esse padrão para o Amazon EMR deixa mais espaço nos contêineres do YARN para a sobrecarga de memória do executor do que o padrão de 10% definido internamente pelo Spark. O padrão do Amazon EMR de 18,75% mostrou empiricamente menos falhas relacionadas à memória nas avaliações comparativas do TPC-DS.
  - Backport do SPARK-26316 para melhorar a performance.
- Nas versões 5.19.0, 5.20.0 e 5.21.0 do Amazon EMR, os rótulos dos nós do YARN são armazenados em um diretório do HDFS. Em algumas situações, isso leva a atrasos na inicialização do nó central causando, em seguida, tempo limite do cluster e falha na inicialização.

Desde a versão 5.22.0 do Amazon EMR, esse problema foi resolvido. Os rótulos dos nós do YARN são armazenados no disco local de cada nó do cluster, evitando dependências no HDFS.

Problemas conhecidos

- Hue (corrigido na versão 5.24.0 do Amazon EMR)
  - O Hue executado no Amazon EMR não é compatível com o Solr. Desde a versão 5.20.0 do Amazon EMR, um problema de configuração incorreta faz com que o Solr seja habilitado e uma mensagem de erro inofensiva semelhante à seguinte seja exibida:

    ```
    Solr server could not be contacted properly:
    HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
    port=1978): Max retries exceeded with url: /solr/admin/info/
    system?user.name=hue&doAs=administrator&wt=json (Caused by
    NewConnectionError(': Failed to establish a new connection: [Errno 111]
    Connection refused',))
    ```

    Para evitar que a mensagem de erro do Solr seja exibida:

    1. Conecte-se à linha de comando do nó primário usando SSH.
    2. Use um editor de texto para abrir o arquivo `hue.ini`. Por exemplo:

       ```
       sudo vim /etc/hue/conf/hue.ini
       ```

    3. Pesquise o termo `appblacklist` e modifique a linha para o seguinte:

       ```
       appblacklist = search
       ```

    4. Salve as alterações e reinicie o Hue, conforme mostrado no exemplo a seguir:

       ```
       sudo stop hue; sudo start hue
       ```

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução

automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

Solução:

- SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
- Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versão 5.21.1

As notas da versão a seguir incluem informações para a versão 5.21.1 do Amazon EMR. As alterações são referentes à versão 5.21.0.

Data da versão inicial: 18 de julho de 2019

Alterações, melhorias e problemas resolvidos

- Atualizada a AMI padrão do Amazon Linux para Amazon EMR para incluir atualizações de segurança importantes do kernel Linux, incluindo o problema de negação de serviço do TCP SACK (AWS-2019-005).

Problemas conhecidos

- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

Solução:

- SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
- Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versão 5.21.0

As notas da versão a seguir incluem informações para a versão 5.21.0 do Amazon EMR. As alterações são referentes à versão 5.20.0.

Data da versão inicial: 18 de fevereiro de 2019

Data da última atualização: 3 de abril de 2019

Atualizações

- Flink 1.7.0

- Presto 0.215
- AWS SDK para Java 1.11.479

Novos recursos

- (3 de abril de 2019) Com as versões 5.21.0 e posteriores do Amazon EMR, você pode substituir as configurações de cluster e especificar classificações de configuração adicionais para cada grupo de instâncias em um cluster em execução. Você faz isso usando o console do Amazon EMR, o AWS Command Line Interface (AWS CLI) ou o AWS SDK. Para obter mais informações, consulte Supplying a Configuration for an Instance Group in a Running Cluster.

Alterações, melhorias e problemas resolvidos

- Zeppelin
  - Backport do ZEPPELIN-3878.

Problemas conhecidos

- Hue (corrigido na versão 5.24.0 do Amazon EMR)
  - O Hue executado no Amazon EMR não é compatível com o Solr. Desde a versão 5.20.0 do Amazon EMR, um problema de configuração incorreta faz com que o Solr seja habilitado e uma mensagem de erro inofensiva semelhante à seguinte seja exibida:

    ```
    Solr server could not be contacted properly:
    HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
    port=1978): Max retries exceeded with url: /solr/admin/info/
    system?user.name=hue&doAs=administrator&wt=json (Caused by
    NewConnectionError(': Failed to establish a new connection: [Errno 111]
    Connection refused',))
    ```

    Para evitar que a mensagem de erro do Solr seja exibida:

    1. Conecte-se à linha de comando do nó primário usando SSH.
    2. Use um editor de texto para abrir o arquivo `hue.ini`. Por exemplo:

       ```
       sudo vim /etc/hue/conf/hue.ini
       ```

    3. Pesquise o termo `appblacklist` e modifique a linha para o seguinte:

```
appblacklist = search
```

4. Salve as alterações e reinicie o Hue, conforme mostrado no exemplo a seguir:

```
sudo stop hue; sudo start hue
```

- Tez
  - Esse problema foi corrigido no Amazon EMR 5.22.0.

    Quando você se conecta à interface do usuário do Tez em http: //:8080/tez-ui *MasterDNS* por meio de uma conexão SSH com o nó primário do cluster, aparece o erro “Falha na operação do adaptador - o servidor Timeline (ATS) está fora de alcance. Ele está inoperante ou o CORS não está habilitado” ou as tarefas mostram, inesperadamente, N/A.

    Isso é causado pela interface do usuário do Tez ao fazer solicitações ao servidor de linha do tempo do YARN usando `localhost` em vez do nome do host do nó primário. Como solução alternativa, um script está disponível para execução como ação ou etapa de bootstrap. O script atualiza o nome do host no arquivo `configs.env` do Tez. Para obter mais informações e a localização do script, consulte Instruções de bootstrap.
- Nas versões 5.19.0, 5.20.0 e 5.21.0 do Amazon EMR, os rótulos dos nós do YARN são armazenados em um diretório do HDFS. Em algumas situações, isso leva a atrasos na inicialização do nó central causando, em seguida, tempo limite do cluster e falha na inicialização. Desde a versão 5.22.0 do Amazon EMR, esse problema foi resolvido. Os rótulos dos nós do YARN são armazenados no disco local de cada nó do cluster, evitando dependências no HDFS.
- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

  Solução:

  - SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.

- Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do hadoop.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, o arquivo keytab está localizado em /etc/hadoop.keytab e a entidade principal está na forma de hadoop/<hostname>@<REALM>.

> **Note**
>
> Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versão 5.20.0

As notas da versão a seguir incluem informações para a versão 5.20.0 do Amazon EMR. As alterações são referentes à versão 5.19.0.

Data da versão inicial: 18 de dezembro de 2018

Data da última atualização: 22 de janeiro de 2019

Atualizações

- Flink 1.6.2
- HBase 1.4.8
- Hive 2.3.4
- Hue 4.3.0
- MXNet 1.3.1
- Presto 0.214
- Spark 2.4.0
- TensorFlow 1.12.0
- Tez 0.9.1
- AWS SDK para Java 1.11.461

Novos recursos

- (22 de janeiro de 2019) O Kerberos no Amazon EMR foi aprimorado para oferecer suporte à autenticação de entidades principais de um KDC externo. Isso centraliza o gerenciamento de principais porque vários clusters podem compartilhar um único KDC externo. Além disso, o KDC externo pode ter uma relação de confiança entre realm com um domínio do Active Directory. Isso permite que todos os clusters autentiquem principais do Active Directory. Para obter mais informações, consulte Usar autenticação Kerberos no Guia de gerenciamento do Amazon EMR.

Alterações, melhorias e problemas resolvidos

- AMI padrão do Amazon Linux para Amazon EMR
  - O pacote Python3 foi atualizado do python 3.4 para 3.6.
- O confirmador otimizado para EMRFS S3
  - O confirmador otimizado para EMRFS S3 agora está habilitado por padrão, o que melhora a performance de gravação. Para obter mais informações, consulte Usar o confirmador otimizado para EMRFS S3.
- Hive
  - Backport do HIVE-16686.
- Glue com Spark e Hive
  - No EMR 5.20.0 ou posterior, a remoção paralela de partições é ativada automaticamente para Spark e Hive quando o AWS Glue Data Catalog é usado como metastore. Essa alteração reduz significativamente o tempo de planejamento de consultas ao executar várias solicitações em paralelo para recuperar partições. O número total de segmentos que podem ser executados simultaneamente varia entre 1 e 10. O valor padrão é 5, que é uma configuração recomendada. Você pode alterá-lo especificando a propriedade `aws.glue.partition.num.segments` na classificação de configuração `hive-site`. Se ocorrer controle de utilização, você poderá desativar o atributo alterando o valor para 1. Para obter mais informações, consulte a Estrutura de segmentos do AWS Glue.

Problemas conhecidos

- Hue (corrigido na versão 5.24.0 do Amazon EMR)

- O Hue executado no Amazon EMR não é compatível com o Solr. Desde a versão 5.20.0 do Amazon EMR, um problema de configuração incorreta faz com que o Solr seja habilitado e uma mensagem de erro inofensiva semelhante à seguinte seja exibida:

  ```
  Solr server could not be contacted properly:
  HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
  port=1978): Max retries exceeded with url: /solr/admin/info/
  system?user.name=hue&doAs=administrator&wt=json (Caused by
  NewConnectionError(': Failed to establish a new connection: [Errno 111]
  Connection refused',))
  ```

  Para evitar que a mensagem de erro do Solr seja exibida:

  1. Conecte-se à linha de comando do nó primário usando SSH.
  2. Use um editor de texto para abrir o arquivo `hue.ini`. Por exemplo:

     ```
     sudo vim /etc/hue/conf/hue.ini
     ```
  3. Pesquise o termo `appblacklist` e modifique a linha para o seguinte:

     ```
     appblacklist = search
     ```
  4. Salve as alterações e reinicie o Hue, conforme mostrado no exemplo a seguir:

     ```
     sudo stop hue; sudo start hue
     ```

- Tez
  - Esse problema foi corrigido no Amazon EMR 5.22.0.

  Quando você se conecta à interface do usuário do Tez em http: //:8080/tez-ui *MasterDNS* por meio de uma conexão SSH com o nó primário do cluster, aparece o erro “Falha na operação do adaptador - o servidor Timeline (ATS) está fora de alcance. Ele está inoperante ou o CORS não está habilitado” ou as tarefas mostram, inesperadamente, N/A.

  Isso é causado pela interface do usuário do Tez ao fazer solicitações ao servidor de linha do tempo do YARN usando `localhost` em vez do nome do host do nó primário. Como solução alternativa, um script está disponível para execução como ação ou etapa de bootstrap. O script atualiza o nome do host no arquivo `configs.env` do Tez. Para obter mais informações e a localização do script, consulte Instruções de bootstrap.

- Nas versões 5.19.0, 5.20.0 e 5.21.0 do Amazon EMR, os rótulos dos nós do YARN são armazenados em um diretório do HDFS. Em algumas situações, isso leva a atrasos na inicialização do nó central causando, em seguida, tempo limite do cluster e falha na inicialização. Desde a versão 5.22.0 do Amazon EMR, esse problema foi resolvido. Os rótulos dos nós do YARN são armazenados no disco local de cada nó do cluster, evitando dependências no HDFS.
- Problema conhecido em clusters com vários nós primários e autenticação Kerberos

  Se você executar clusters com vários nós primários e autenticação Kerberos nas versões 5.20.0 e posteriores do Amazon EMR, poderá encontrar problemas nas operações de cluster, como redução da escala verticalmente ou envio de etapas depois que o cluster estiver em execução por algum tempo. O período depende do período de validade do tíquete do Kerberos que você definiu. O problema de redução da escala verticalmente afeta tanto as solicitações de redução automática quanto as de reduções explícitas que você enviou. Operações adicionais de cluster também podem ser afetadas.

  Solução:

  - SSH como usuário do `hadoop` para o nó primário de liderança do cluster do EMR com vários nós primários.
  - Execute o comando a seguir para renovar o tíquete do Kerberos para o usuário do `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, o arquivo keytab está localizado em `/etc/hadoop.keytab` e a entidade principal está na forma de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Essa solução alternativa entrará em vigor durante o período de validade do tíquete do Kerberos. Essa duração é de 10 horas por padrão, mas pode ser configurada pelas definições do Kerberos. Você deve executar novamente o comando acima quando o tíquete do Kerberos expirar.

Versão 5.19.0

As notas da versão a seguir incluem informações para a versão 5.19.0 do Amazon EMR. As alterações são referentes à versão 5.18.0.

Data da versão inicial: 7 de novembro de 2018

Data da última atualização: 19 de novembro de 2018

Atualizações

- Hadoop 2.8.5
- Flink 1.6.1
- JupyterHub 0.9.4
- MXNet 1.3.0
- Presto 0.212
- TensorFlow 1.11.0
- Zookeeper 3.4.13
- AWS SDK para Java 1.11.433

Novos recursos

- (19 de novembro de 2018) Os Cadernos do EMR constituem um ambiente gerenciado baseado no caderno Jupyter. Ele suporta os kernels mágicos do Spark para Spark PySpark SQL, Spark R e Scala. Os Cadernos do EMR podem ser usados com clusters criados usando as versões 5.18.0 e posteriores do Amazon EMR. Para obter mais informações, consulte Usar Cadernos do EMR no Guia de gerenciamento do Amazon EMR.
- O confirmador otimizado para EMRFS S3 está disponível ao serem gravados arquivos Parquet usando Spark e EMRFS. Esse confirmador melhora a performance de gravação. Para obter mais informações, consulte Usar o confirmador otimizado para EMRFS S3.

Alterações, melhorias e problemas resolvidos

- YARN
  - Modificada a lógica que limita o processo mestre da aplicação à execução nos nós centrais. Essa funcionalidade agora usa o atributo e as propriedades de rótulos de nós do YARN nas classificações de configuração `yarn-site` e `capacity-scheduler`. Para obter mais informações, consulte https://docs.aws.amazon.com/emr/latest/ManagementGuide/emr-plan-instances-guidelines.html#emr-plan-spot-YARN.
- AMI padrão do Amazon Linux para Amazon EMR

- `ruby18`, `php56`, e `gcc48` não são mais instalados por padrão. Eles podem ser instalados, se desejado, usando `yum`.
- A gem do ruby aws-sdk não é mais instalada por padrão. Ela pode ser instalada usando `gem install aws-sdk`, se desejado. Componentes específicos também podem ser instalados. Por exemplo, `.gem install aws-sdk-s3`

Problemas conhecidos

- Cadernos do EMR: em algumas circunstâncias, com vários editores de cadernos abertos, o editor de cadernos pode parecer incapaz de se conectar ao cluster. Se isso acontecer, limpe os cookies do navegador e reabra os editores de cadernos.
- CloudWatch ContainerPending Escalabilidade métrica e automática — (corrigida na versão 5.20.0) O Amazon EMR pode emitir um valor negativo para. `ContainerPending` Se `ContainerPending` for usado em uma regra de escalabilidade automática, a escalabilidade automática não se comportará conforme esperado. Evite usar `ContainerPending` com escalabilidade automática.
- Nas versões 5.19.0, 5.20.0 e 5.21.0 do Amazon EMR, os rótulos dos nós do YARN são armazenados em um diretório do HDFS. Em algumas situações, isso leva a atrasos na inicialização do nó central causando, em seguida, tempo limite do cluster e falha na inicialização. Desde a versão 5.22.0 do Amazon EMR, esse problema foi resolvido. Os rótulos dos nós do YARN são armazenados no disco local de cada nó do cluster, evitando dependências no HDFS.

Versão 5.18.0

As notas da versão a seguir incluem informações para a versão 5.18.0 do Amazon EMR. As alterações são referentes à versão 5.17.0.

Data da versão inicial: 24 de outubro de 2018

Atualizações

- Flink 1.6.0
- HBase 1.4.7
- Presto 0.210
- Spark 2.3.2
- Zeppelin 0.8.0

Novos recursos

- Desde a versão 5.18.0 do Amazon EMR, você pode usar o repositório de artefatos do Amazon EMR para criar o código de trabalho em comparação com as versões exatas de bibliotecas e dependências disponíveis com versões específicas do Amazon EMR. Para obter mais informações, consulte Verificar dependências usando o repositório de artefatos do Amazon EMR.

Alterações, melhorias e problemas resolvidos

- Hive
  - Adicionado suporte para o S3 Select. Para obter mais informações, consulte Usar o S3 Select com o Hive para melhorar a performance.
- Presto
  - Adicionado suporte para o S3 Select Pushdown. Para obter mais informações, consulte Usar S3 Select Pushdown com o Presto para melhorar a performance.
- Spark
  - A configuração log4j padrão do Spark foi alterada para lançar logs de contêineres por hora para trabalhos de streaming do Spark. Isso ajuda a evitar a exclusão de lo de trabalhos de streaming do Spark de execução prolongada.

Versão 5.17.1

As notas da versão a seguir incluem informações para a versão 5.17.1 do Amazon EMR. As alterações são referentes à versão 5.17.0.

Data da versão inicial: 18 de julho de 2019

Alterações, melhorias e problemas resolvidos

- Atualizada a AMI padrão do Amazon Linux para Amazon EMR para incluir atualizações de segurança importantes do kernel Linux, incluindo o problema de negação de serviço do TCP SACK (AWS-2019-005).

Versão 5.17.0

As notas da versão a seguir incluem informações para a versão 5.17.0 do Amazon EMR. As alterações são referentes à versão 5.16.0.

Data da versão inicial: 30 de agosto de 2018

Atualizações

- Flink 1.5.2
- HBase 1.4.6
- Presto 0.206

Novos recursos

- Adicionado suporte para Tensorflow. Para obter mais informações, consulte TensorFlow.

Alterações, melhorias e problemas resolvidos

- JupyterHub
  - Adicionado suporte para a persistência de cadernos no Amazon S3. Para obter mais informações, consulte Configurar a persistência de cadernos no Amazon S3.
- Spark
  - Adicionado suporte para o S3 Select. Para obter mais informações, consulte Usar o S3 Select com Spark para melhorar a performance das consultas.
- Resolvidos os problemas com as métricas do Cloudwatch e o atributo de escalabilidade automática nas versões 5.14.0, 5.15.0 ou 5.16.0 do Amazon EMR.

Problemas conhecidos

- Quando você cria um cluster kerberizado com o Livy instalado, o Livy apresenta falha com um erro em que a autenticação simples não está habilitada. A reinicialização do servidor do Livy resolve o problema. Como solução alternativa, adicione uma etapa durante a criação do cluster que execute `sudo restart livy-server` no nó primário.
- Se você usar uma AMI do Amazon Linux personalizada com base em uma AMI do Amazon Linux com data de criação 11/8/2018, o servidor Oozie falhará ao iniciar. Se você usar o Oozie, crie uma AMI personalizada com base em um ID de AMI do Amazon Linux com uma data de criação diferente. Você pode usar o AWS CLI comando a seguir para retornar uma lista de imagens IDs para todos os HVM Amazon Linux AMIs com uma versão 2018.03, junto com a data de lançamento, para que você possa escolher uma Amazon Linux AMI apropriada como sua base. MyRegion Substitua pelo seu identificador de região, como us-west-2.

```
aws ec2 --region MyRegion describe-images --owner amazon --query 'Images[?
Name!=`null`]|[?starts_with(Name, `amzn-ami-hvm-2018.03`) == `true`].
[CreationDate,ImageId,Name]' --output text | sort -rk1
```

Versão 5.16.0

As notas da versão a seguir incluem informações para a versão 5.16.0 do Amazon EMR. As alterações são referentes à versão 5.15.0.

Data da versão inicial: 19 de julho de 2018

Atualizações

- Hadoop 2.8.4
- Flink 1.5.0
- Livy 0.5.0
- MXNet 1.2.0
- Phoenix 4.14.0
- Presto 0.203
- Spark 2.3.1
- AWS SDK para Java 1.11.336
- CUDA 9.2
- Driver JDBC do Redshift, 1.2.15.1025

Alterações, melhorias e problemas resolvidos

- HBase
  - Backport do HBASE-20723
- Presto
  - Alterações na configuração para oferecer suporte à autenticação LDAP. Para obter mais informações, consulte Usar autenticação LDAP para o Presto no Amazon EMR.
- Spark

- A versão 2.3.1 do Apache Spark, disponível desde a versão 5.16.0 do Amazon EMR, aborda CVE-2018-8024 e CVE-2018-1334. Recomendamos que você migre as versões anteriores do Spark para a versão 2.3.1 ou posteriores.

Problemas conhecidos

- Essa versão não é compatível com os tipos de instância c1.medium ou m1.small. Os clusters que usam qualquer um desses tipos de instância não são iniciados. Como solução alternativa, especifique um tipo de instância diferente ou use uma versão diferente.
- Quando você cria um cluster kerberizado com o Livy instalado, o Livy apresenta falha com um erro em que a autenticação simples não está habilitada. A reinicialização do servidor do Livy resolve o problema. Como solução alternativa, adicione uma etapa durante a criação do cluster que execute `sudo restart livy-server` no nó primário.
- Depois que o nó primário for reinicializado ou o controlador de instância for reiniciado, as CloudWatch métricas não serão coletadas e o recurso de escalabilidade automática não estará disponível nas versões 5.14.0, 5.15.0 ou 5.16.0 do Amazon EMR. Esse problema foi corrigido na versão 5.17.0 do Amazon EMR.

Versão 5.15.0

As notas da versão a seguir incluem informações para a versão 5.15.0 do Amazon EMR. As alterações são referentes à versão 5.14.0.

Data da versão inicial: 21 de junho de 2018

Atualizações

- Atualizado HBase para 1.4.4
- Atualizado Hive para 2.3.3
- Atualizado Hue para 4.2.0
- Atualizado Oozie para 5.0.0
- Atualizado Zookeeper para 3.4.12
- AWS SDK atualizado para 1.11.333

Alterações, melhorias e problemas resolvidos

- Hive
  - Backport do HIVE-18069
- Hue
  - Atualizado o Hue para se autenticar corretamente com o Livy quando o Kerberos está habilitado. Agora, o Livy é compatível quando usa o Kerberos com o Amazon EMR.
- JupyterHub
  - Atualizado JupyterHub para que o Amazon EMR instale bibliotecas de clientes LDAP por padrão.
  - Corrigido um erro no script que gera certificados autoassinados.

Problemas conhecidos

- Essa versão não é compatível com os tipos de instância c1.medium ou m1.small. Os clusters que usam qualquer um desses tipos de instância não são iniciados. Como solução alternativa, especifique um tipo de instância diferente ou use uma versão diferente.
- Depois que o nó primário for reinicializado ou o controlador de instância for reiniciado, as CloudWatch métricas não serão coletadas e o recurso de escalabilidade automática não estará disponível nas versões 5.14.0, 5.15.0 ou 5.16.0 do Amazon EMR. Esse problema foi corrigido na versão 5.17.0 do Amazon EMR.

Versão 5.14.1

As notas da versão a seguir incluem informações para a versão 5.14.1 do Amazon EMR. As alterações são referentes à versão 5.14.0.

Data da versão inicial: 17 de outubro de 2018

Atualizada a AMI padrão para Amazon EMR para abordar possíveis vulnerabilidades de segurança.

Versão 5.14.0

As notas da versão a seguir incluem informações para a versão 5.14.0 do Amazon EMR. As alterações são referentes à versão 5.13.0.

Data da versão inicial: 4 de junho de 2018

Atualizações

- Atualizado Apache Flink para 1.4.2
- Apache MXnet atualizado para 1.1.0
- Atualizado Apache Sqoop para 1.4.7

Novos recursos

- JupyterHub Suporte adicionado. Para obter mais informações, consulte JupyterHub.

Alterações, melhorias e problemas resolvidos

- EMRFS
  - A string userAgent nas solicitações ao Amazon S3 foi atualizada para conter as informações de usuário e grupo da entidade principal invocadora. Isso pode ser usado com AWS CloudTrail registros para um rastreamento de solicitações mais abrangente.
- HBase
  - Incluído o HBASE-20447, que aborda um problema que poderia causar falhas de cache, especialmente com regiões divididas.
- MXnet
  - Adicionadas bibliotecas OpenCV.
- Spark
  - Quando o Spark grava arquivos Parquet em um local do Amazon S3 usando o EMRFS, FileOutputCommitter o algoritmo foi atualizado para usar a versão 2 em vez da versão 1. Isso reduz o número de renomeações, o que melhora a performance da aplicação. Essa alteração não afeta:
    - Aplicações diferentes do Spark.
    - Aplicativos que gravam em outros sistemas de arquivos, como o HDFS (que ainda usa a versão 1 do FileOutputCommitter).
    - Aplicações que usam outros formatos de saída, como texto ou csv, que já usam a gravação direta do EMRFS.

Problemas conhecidos

- JupyterHub
  - O uso de classificações de configuração para configurar JupyterHub notebooks Jupyter individuais ao criar um cluster não é suportado. Edite manualmente o arquivo jupyterhub_config.py e os arquivos jupyter_notebook_config.py para cada usuário. Para obter mais informações, consulte Configurando JupyterHub.
  - JupyterHub falha ao iniciar em clusters dentro de uma sub-rede privada, falhando com a mensagem. `Error: ENOENT: no such file or directory, open '/etc/jupyter/conf/server.crt'` Isso é causado por um erro no script que gera certificados autoassinados. Use a solução alternativa a seguir para gerar certificados autoassinados. Todos os comandos são executados enquanto estão conectados ao nó primário.

    1. Copie o script de geração de certificados do contêiner para o nó primário:

       ```
       sudo docker cp jupyterhub:/tmp/gen_self_signed_cert.sh ./
       ```

    2. Use um editor de texto para alterar a linha 23 e mudar o nome de host público para o nome deo host local, conforme mostrado abaixo:

       ```
       local hostname=$(curl -s $EC2_METADATA_SERVICE_URI/local-hostname)
       ```

    3. Execute o script para gerar certificados autoassinados:

       ```
       sudo bash ./gen_self_signed_cert.sh
       ```

    4. Mova os arquivos de certificado que o script gera para o diretório `/etc/jupyter/conf/`:

       ```
       sudo mv /tmp/server.crt /tmp/server.key /etc/jupyter/conf/
       ```

    Você pode acessar `tail` o `jupyter.log` arquivo para verificar se ele JupyterHub foi reiniciado e está retornando um código de resposta 200. Por exemplo:

    ```
    tail -f /var/log/jupyter/jupyter.log
    ```

    Essa ação deve retornar uma resposta semelhante à seguinte:

    ```
    # [I 2018-06-14 18:56:51.356 JupyterHub app:1581] JupyterHub is now running at
     https://:9443/
    ```

```
# 19:01:51.359 - info: [ConfigProxy] 200 GET /api/routes
```

- Depois que o nó primário for reinicializado ou o controlador de instância for reiniciado, as CloudWatch métricas não serão coletadas e o recurso de escalabilidade automática não estará disponível nas versões 5.14.0, 5.15.0 ou 5.16.0 do Amazon EMR. Esse problema foi corrigido na versão 5.17.0 do Amazon EMR.

Versão 5.13.0

As notas da versão a seguir incluem informações para a versão 5.13.0 do Amazon EMR. As alterações são referentes à versão 5.12.0.

Atualizações

- Atualizado Spark para 2.3.0
- Atualizado HBase para 1.4.2
- Atualizado Presto para 0.194
- Atualizado para AWS SDK para Java 1.11.297

Alterações, melhorias e problemas resolvidos

- Hive
  - Backport do HIVE-15436. Hive aprimorado APIs para retornar somente visualizações.

Problemas conhecidos

- MXNet atualmente não tem bibliotecas OpenCV.

Versão 5.12.2

As notas da versão a seguir incluem informações para a versão 5.12.2 do Amazon EMR. As alterações são referentes à versão 5.12.1.

Data da versão inicial: 29 de agosto de 2018

Alterações, melhorias e problemas resolvidos

- Esta versão aborda uma possível vulnerabilidade de segurança.

Versão 5.12.1

As notas da versão a seguir incluem informações para a versão 5.12.1 do Amazon EMR. As alterações são referentes à versão 5.12.0.

Data da versão inicial: 29 de março de 2018

Alterações, melhorias e problemas resolvidos

- Atualizado o kernel do Amazon Linux da AMI padrão do Amazon Linux para Amazon EMR para abordar possíveis vulnerabilidades.

Versão 5.12.0

As notas da versão a seguir incluem informações para a versão 5.12.0 do Amazon EMR. As alterações são referentes à versão 5.11.1.

Atualizações

- AWS SDK para Java 1.11.238 ⇒ 1.11.267. Para obter mais informações, consulte o AWS SDK for Java Change Log on GitHub.
- Hadoop 2.7.3 ⇒ 2.8.3. Para obter mais informações, consulte Versões do Apache Hadoop.
- Flink 1.3.2 ⇒ 1.4.0. Para obter mais informações, consulte o Anúncio de versão do Apache Flink 1.4.0.
- HBase 1.3.1 ⇒ 1.4.0. Para obter mais informações, consulte o anúncio HBase de lançamento.
- Hue 4.0.1 ⇒ 4.1.0. Para obter mais informações, consulte as Notas de versão.
- MxNet 0,12,0 ⇒ 1,0,0. Para obter mais informações, consulte o MXNet Change Log on GitHub.
- Presto 0.187 ⇒ 0.188. Para obter mais informações, consulte as Notas de versão.

Alterações, melhorias e problemas resolvidos

- Hadoop
  - A propriedade `yarn.resourcemanager.decommissioning.timeout` foi alterada para `yarn.resourcemanager.nodemanager-graceful-decommission-timeout-secs`. Você pode usar essa propriedade para personalizar a redução da escala do cluster verticalmente. Para obter mais informações, consulte Redução da escala do cluster verticalmente no Guia de gerenciamento do Amazon EMR.

  - A CLI do Hadoop adicionou a opção `-d` ao comando `cp` (copy), que especifica cópia direta. Você pode usar isso para evitar a criação de um arquivo `.COPYING` intermediário, o que torna mais rápida a cópia de dados entre o Amazon S3. Para obter mais informações, consulte HADOOP-12384.
- Pig
  - Adicionada a classificação de configuração `pig-env`, que simplifica a configuração das propriedades do ambiente do Pig. Para obter mais informações, consulte Configurar aplicações.
- Presto
  - Adicionada a classificação de configuração `presto-connector-redshift`, que pode ser usada para configurar os valores no arquivo de configuração `redshift.properties` do Presto. Para obter mais informações, consulte Conector do Redshift na documentação do Presto e Configurar aplicações.
  - O suporte do Presto para EMRFS foi adicionado e é a configuração padrão. As versões anteriores do Amazon EMR usavam o PrestOS3FileSystem, que era a única opção. Para obter mais informações, consulte Configuração do EMRFS e do PrestoS3 FileSystem.

    > **Note**
    >
    > Se você consultar dados subjacentes no Amazon S3 com a versão 5.12.0 do Amazon EMR, poderão ocorrer erros no Presto. Isso acontece porque o Presto não consegue obter valores de classificação de configuração em `emrfs-site.xml`. Como solução alternativa, crie um subdiretório `emrfs` em `usr/lib/presto/plugin/hive-hadoop2/` e crie um link simbólico em `usr/lib/presto/plugin/hive-hadoop2/emrfs` para o arquivo `/usr/share/aws/emr/emrfs/conf/emrfs-site.xml` existente. Em seguida, reinicie o processo presto-server (`sudo presto-server stop` seguido por `sudo presto-server start`).
- Spark
  - SPARK-22036 retroportado: a BigDecimal multiplicação às vezes retorna nula.

Problemas conhecidos

- MXNet não inclui bibliotecas OpenCV.
- O SparkR não está disponível para clusters criados usando uma AMI personalizada porque R não é instalado por padrão nos nós do cluster.

Versão 5.11.3

As notas da versão a seguir incluem informações para a versão 5.11.3 do Amazon EMR. As alterações são referentes à versão 5.11.2.

Data da versão inicial: 18 de julho de 2019

Alterações, melhorias e problemas resolvidos

- Atualizada a AMI padrão do Amazon Linux para Amazon EMR para incluir atualizações de segurança importantes do kernel Linux, incluindo o problema de negação de serviço do TCP SACK (AWS-2019-005).

Versão 5.11.2

As notas da versão a seguir incluem informações para a versão 5.11.2 do Amazon EMR. As alterações são referentes à versão 5.11.1.

Data da versão inicial: 29 de agosto de 2018

Alterações, melhorias e problemas resolvidos

- Esta versão aborda uma possível vulnerabilidade de segurança.

Versão 5.11.1

As notas da versão a seguir incluem informações para a versão 5.11.1 do Amazon EMR. As alterações são referentes à versão 5.11.0 do Amazon EMR.

Data da versão inicial: 22 de janeiro de 2018

Alterações, melhorias e problemas resolvidos

- Atualizado o kernel do Amazon Linux da AMI padrão do Amazon Linux para Amazon EMR para abordar vulnerabilidades associadas à execução especulativa (CVE-2017-5715, CVE-2017-5753 e CVE-2017-5754). Para obter mais informações, consulte https://aws.amazon.com/security/security-bulletins/AWS-2018-013/.

Problemas conhecidos

- MXNet não inclui bibliotecas OpenCV.
- Por padrão, o Hive 2.3.2 define `hive.compute.query.using.stats=true`. Isso faz com que as consultas obtenham dados de estatísticas existentes em vez de diretamente dos dados, o que pode gerar confusão. Por exemplo, se você tiver uma tabela com `hive.compute.query.using.stats=true` e fizer upload de novos arquivos para a tabela `LOCATION`, a execução de uma consulta `SELECT COUNT(*)` na tabela retornará a contagem das estatísticas, e não selecionará as linhas adicionadas.

  Como alternativa, use o comando `ANALYZE TABLE` para reunir novas estatísticas ou defina `hive.compute.query.using.stats=false`. Para obter mais informações, consulte Estatísticas no Hive na documentação do Apache Hive.

Versão 5.11.0

As notas da versão a seguir incluem informações para a versão 5.11.0 do Amazon EMR. As alterações são referentes à versão 5.10.0 do Amazon EMR.

Atualizações

Os aplicativos e os componentes a seguir foram atualizados nesta versão para incluir as seguintes versões.

- Hive 2.3.2
- Spark 2.2.1
- SDK para Java 1.11.238

Novos recursos

- Spark
  - Adicionada a configuração `spark.decommissioning.timeout.threshold`, que aprimora o comportamento de desativação do Spark ao usar instâncias spot. Para obter mais informações, consulte Configurar o comportamento de desativação de nós.
  - Foi adicionado o `aws-sagemaker-spark-sdk` componente ao Spark, que instala o Amazon SageMaker Spark e as dependências associadas para a integração do Spark com a Amazon. SageMaker Você pode usar o Amazon SageMaker Spark para criar pipelines de aprendizado

de máquina (ML) do Spark usando os estágios da Amazon. SageMaker Para obter mais informações, consulte o readme do SageMaker Spark sobre GitHub e como usar o Apache Spark com a Amazon SageMaker no Amazon Developer Guide. SageMaker

Problemas conhecidos

- MXNet não inclui bibliotecas OpenCV.
- Por padrão, o Hive 2.3.2 define `hive.compute.query.using.stats=true`. Isso faz com que as consultas obtenham dados de estatísticas existentes em vez de diretamente dos dados, o que pode gerar confusão. Por exemplo, se você tiver uma tabela com `hive.compute.query.using.stats=true` e fizer upload de novos arquivos para a tabela `LOCATION`, a execução de uma consulta `SELECT COUNT(*)` na tabela retornará a contagem das estatísticas, e não selecionará as linhas adicionadas.

  Como alternativa, use o comando `ANALYZE TABLE` para reunir novas estatísticas ou defina `hive.compute.query.using.stats=false`. Para obter mais informações, consulte Estatísticas no Hive na documentação do Apache Hive.

Versão 5.10.0

As notas da versão a seguir incluem informações para a versão 5.10.0 do Amazon EMR. As alterações são referentes à versão 5.9.0 do Amazon EMR.

Atualizações

Os aplicativos e os componentes a seguir foram atualizados nesta versão para incluir as seguintes versões.

- AWS SDK para Java 1.11.221
- Hive 2.3.1
- Presto 0.187

Novos recursos

- Adicionado o suporte para autenticação do Kerberos. Para obter mais informações, consulte Usar autenticação Kerberos no Guia de gerenciamento do Amazon EMR

- Adicionado suporte para perfis do IAM para solicitações do EMRFS para o Amazon S3. Para obter mais informações, consulte Configurar perfis do IAM para solicitações do EMRFS para o Amazon S3 no Guia de gerenciamento do Amazon EMR.
- Suporte adicionado para tipos de instância P2 e P3 baseados em GPU. Para obter mais informações, consulte Instâncias Amazon EC2 P2 e Instâncias Amazon EC2 P3. Os drivers NVIDIA 384.81 e CUDA driver 9.0.176 são instalados nesses tipos de instância por padrão.
- O suporte adicionado para Apache MXNet.

Alterações, melhorias e problemas resolvidos

- Presto
  - Foi adicionado suporte para usar o AWS Glue Data Catalog como metastore padrão do Hive. Para obter mais informações, consulte Usando o Presto com o AWS Glue Data Catalog.
  - Suporte adicionado para funções geoespaciais.
  - Suporte de vazamento para disco adicionado para uniões.
  - Suporte adicionado para o Conector Redshift.
- Spark
  - Backport SPARK-20640, que causa tempo limite rpc e as repetições de valores de registro configuráveis usando as propriedades `spark.shuffle.registration.timeout` e `spark.shuffle.registration.maxAttempts`.
  - SPARK-21549 retroportado, que corrige um erro que ocorre ao gravar de forma personalizada em locais que não sejam do HDFS. OutputFormat
- Backport Hadoop-13270
- As bibliotecas Numpy, Scipy e Matplotlib foram removidas da AMI base do Amazon EMR. Se forem necessárias para o aplicativo, essas bibliotecas estarão disponíveis no repositório do aplicativo. Portanto, você pode usar uma ação de bootstrap para instalá-las em todos os nós usando `yum install`.
- A AMI base do Amazon EMR não tem mais pacotes do RPM de aplicações incluídos, de maneira que os pacotes do RPM não estão mais presentes em nós de cluster. A AMI personalizada AMIs e a AMI básica do Amazon EMR agora fazem referência ao repositório de pacotes RPM no Amazon S3.
- Devido à introdução do faturamento por segundo na Amazon EC2, o comportamento padrão de redução de escala agora é Encerrar na conclusão da tarefa, em vez de Encerrar na hora da

instância. Para obter mais informações, consulte Configurar redução da escala verticalmente do cluster.

Problemas conhecidos

- MXNet não inclui bibliotecas OpenCV.
- Por padrão, o Hive 2.3.1 define `hive.compute.query.using.stats=true`. Isso faz com que as consultas obtenham dados de estatísticas existentes em vez de diretamente dos dados, o que pode gerar confusão. Por exemplo, se você tiver uma tabela com `hive.compute.query.using.stats=true` e fizer upload de novos arquivos para a tabela `LOCATION`, a execução de uma consulta `SELECT COUNT(*)` na tabela retornará a contagem das estatísticas, e não selecionará as linhas adicionadas.

  Como alternativa, use o comando `ANALYZE TABLE` para reunir novas estatísticas ou defina `hive.compute.query.using.stats=false`. Para obter mais informações, consulte Estatísticas no Hive na documentação do Apache Hive.

Versão 5.9.0

As notas da versão a seguir incluem informações para a versão 5.9.0 do Amazon EMR. As alterações são referentes à versão 5.8.0 do Amazon EMR.

Data do release: 5 de outubro de 2017

Última atualização de recursos: 12 de outubro de 2017

Atualizações

Os aplicativos e os componentes a seguir foram atualizados nesta versão para incluir as seguintes versões.

- AWS SDK para Java versão 1.11.183
- Flink 1.3.2
- Hue 4.0.1
- Pig 0.17.0
- Presto 0.184

Novos recursos

- Adição do suporte ao Livy (versão Livy 0.4.0 - em incubação). Para obter mais informações, consulte Apache Livy.
- Adição de suporte para Hue Notebook para Spark.
- Suporte adicionado para EC2 instâncias Amazon da série i3 (12 de outubro de 2017).

Alterações, melhorias e problemas resolvidos

- Spark
  - Adição de um novo conjunto de recursos que ajudam a garantir que o Spark lide de uma forma mais fácil com o encerramento de nós devido a um redimensionamento manual ou uma solicitação de política de escalabilidade automática. Para obter mais informações, consulte Configurar o comportamento de desativação de nós.
  - O SSL é usado em vez do 3DES para criptografia em trânsito para o serviço de transferência em bloco, o que melhora o desempenho ao usar tipos de EC2 instância da Amazon com AES-NI.
  - SPARK-21494 enviado para backport.
- Zeppelin
  - ZEPPELIN-2377 enviado para backport.
- HBase
  - Foi adicionado o patch HBASE-18533, que permite valores adicionais para HBase BucketCache configuração usando a classificação de configuração. `hbase-site`
- Hue
  - Foi adicionado suporte ao AWS Glue Data Catalog para o editor de consultas Hive no Hue.
  - Por padrão, os superusuários do Hue podem acessar todos os arquivos que os perfis do IAM do Amazon EMR têm permissão para acessar. Os usuários recém-criados não têm automaticamente permissões para acessar o navegador de arquivos Amazon S3 e devem ter as permissões `filebrowser.s3_access` ativadas para o grupo deles.
- Resolvido um problema que fazia com que os dados JSON subjacentes criados com o Catálogo de Dados do AWS Glue ficassem inacessíveis.

Problemas conhecidos

- A inicialização do cluster falha quando todas as aplicações são instaladas e o tamanho padrão do volume raiz do Amazon EBS não é alterado. Como solução alternativa, use o `aws emr create-cluster` comando do AWS CLI e especifique um `--ebs-root-volume-size` parâmetro maior.
- Por padrão, o Hive 2.3.0 define `hive.compute.query.using.stats=true`. Isso faz com que as consultas obtenham dados de estatísticas existentes em vez de diretamente dos dados, o que pode gerar confusão. Por exemplo, se você tiver uma tabela com `hive.compute.query.using.stats=true` e fizer upload de novos arquivos para a tabela `LOCATION`, a execução de uma consulta `SELECT COUNT(*)` na tabela retornará a contagem das estatísticas, e não selecionará as linhas adicionadas.

  Como alternativa, use o comando `ANALYZE TABLE` para reunir novas estatísticas ou defina `hive.compute.query.using.stats=false`. Para obter mais informações, consulte Estatísticas no Hive na documentação do Apache Hive.

Versão 5.8.2

As notas da versão a seguir incluem informações para a versão 5.8.2 do Amazon EMR. As alterações são referentes à versão 5.8.1.

Data da versão inicial: 29 de março de 2018

Alterações, melhorias e problemas resolvidos

- Atualizado o kernel do Amazon Linux da AMI padrão do Amazon Linux para Amazon EMR para abordar possíveis vulnerabilidades.

Versão 5.8.1

As notas da versão a seguir incluem informações para a versão 5.8.1 do Amazon EMR. As alterações são referentes à versão 5.8.0 do Amazon EMR.

Data da versão inicial: 22 de janeiro de 2018

Alterações, melhorias e problemas resolvidos

- Atualizado o kernel do Amazon Linux da AMI padrão do Amazon Linux para Amazon EMR para abordar vulnerabilidades associadas à execução especulativa (CVE-2017-5715, CVE-2017-5753 e

CVE-2017-5754). Para obter mais informações, consulte https://aws.amazon.com/security/security-bulletins/AWS-2018-013/.

Versão 5.8.0

As notas da versão a seguir incluem informações para a versão 5.8.0 do Amazon EMR. As alterações são referentes à versão 5.7.0 do Amazon EMR.

Data da versão inicial: 10 de agosto de 2017

Última atualização de recurso: 25 de setembro de 2017

Atualizações

Os aplicativos e os componentes a seguir foram atualizados nesta versão para incluir as seguintes versões:

- AWS SDK 1.11.160
- Flink 1.3.1
- Hive 2.3.0. Para obter mais informações, consulte Notas de versão no site do Apache Hive.
- Spark 2.2.0. Para obter mais informações, consulte Notas de versão no site do Apache Spark.

Novos recursos

- Adição de suporte para visualização do histórico de aplicativos (25 de setembro de 2017). Para obter mais informações, consulte Visualizar histórico de aplicações no Guia de gerenciamento do Amazon EMR.

Alterações, melhorias e problemas resolvidos

- Integração com o AWS Glue Data Catalog
  - Foi adicionada a capacidade do Hive e do Spark SQL de usar o AWS Glue Data Catalog como armazenamento de metadados do Hive. Para obter mais informações, consulte Usar o Catálogo de Dados do AWS Glue como metastore do Hive e Use o catálogo do AWS Glue Data Catalog com o Spark no Amazon EMR.
- Adicionado o Application history (Histórico do aplicativo) aos detalhes do cluster, o que permite que você visualize dados históricos de aplicativos do YARN e detalhes adicionais de aplicativos

do Spark. Para obter mais informações, consulte Visualizar histórico de aplicações no Guia de gerenciamento do Amazon EMR.

- Oozie
  - OOZIE-2748 enviado para backport.
- Hue
  - HUE-5859 enviado para backport
- HBase
  - Foi adicionado um patch para expor a hora de início do servidor HBase mestre por meio do Java Management Extensions (JMX) usando. `getMasterInitializedTime`
  - Adicionado um patch que melhora a hora de início do cluster.

Problemas conhecidos

- A inicialização do cluster falha quando todas as aplicações são instaladas e o tamanho padrão do volume raiz do Amazon EBS não é alterado. Como solução alternativa, use o `aws emr create-cluster` comando do AWS CLI e especifique um `--ebs-root-volume-size` parâmetro maior.
- Por padrão, o Hive 2.3.0 define `hive.compute.query.using.stats=true`. Isso faz com que as consultas obtenham dados de estatísticas existentes em vez de diretamente dos dados, o que pode gerar confusão. Por exemplo, se você tiver uma tabela com `hive.compute.query.using.stats=true` e fizer upload de novos arquivos para a tabela `LOCATION`, a execução de uma consulta `SELECT COUNT(*)` na tabela retornará a contagem das estatísticas, e não selecionará as linhas adicionadas.

  Como alternativa, use o comando `ANALYZE TABLE` para reunir novas estatísticas ou defina `hive.compute.query.using.stats=false`. Para obter mais informações, consulte Estatísticas no Hive na documentação do Apache Hive.
- Spark: ao usar o Spark, há um problema de vazamento no manipulador de arquivos com o daemon apppusher, o que pode ocorrer para um trabalho de execução prolongada do Spark depois de várias horas ou dias. Para corrigir o problema, conecte-se ao nó principal e digite `sudo /etc/init.d/apppusher stop`. Isso interrompe o daemon apppusher, que o Amazon EMR reiniciará automaticamente.
- Application history
  - Os dados históricos dos executores inativos do Spark não está disponível.

- O histórico do aplicativo não está disponível para clusters que usam uma configuração de segurança para habilitar a criptografia em andamento.

Versão 5.7.0

As notas da versão a seguir incluem informações para a versão 5.7.0 do Amazon EMR. As alterações são referentes à versão 5.6.0 do Amazon EMR.

Data do release: 13 de julho de 2017

Atualizações

- Flink 1.3.0
- Phoenix 4.11.0
- Zeppelin 0.7.2

Novos recursos

- Adicionada a possibilidade de especificar uma AMI do Amazon Linux personalizada ao criar um cluster. Para obter mais informações, consulte Usar uma AMI personalizada.

Alterações, melhorias e problemas resolvidos

- HBase
  - Capacidade adicional para configurar clusters de HBase réplica de leitura. Consulte Usar um cluster de réplica de leitura.
  - Várias correções de erros e melhorias
- Presto: adicionada a capacidade de configurar `node.properties`.
- YARN :adicionada a capacidade de configurar `container-log4j.properties`
- Sqoop: enviado para backport SQOOP-2880, que apresenta um argumento que permite definir o diretório temporário do Sqoop.

Versão 5.6.0

As notas da versão a seguir incluem informações para a versão 5.6.0 do Amazon EMR. As alterações são referentes à versão 5.5.0 do Amazon EMR.

Data do release: 5 de junho de 2017

Atualizações

- Flink 1.2.1
- HBase 1.3.1
- Mahout 0.13.0. Esta é a primeira versão do Mahout compatível com o Spark 2.x nas versões 5.0 e posteriores do Amazon EMR.
- Spark 2.1.1

Alterações, melhorias e problemas resolvidos

- Presto
  - Foi adicionada a capacidade de permitir a comunicação SSL/TLS segura entre os nós do Presto, ativando a criptografia em trânsito usando uma configuração de segurança. Para obter mais informações, consulte Criptografia de dados em trânsito.
  - Presto 7661 enviado para backport, adiciona a opção `VERBOSE` à instrução `EXPLAIN ANALYZE` para relatar estatísticas de baixo nível mais detalhadas sobre um plano de consulta.

Versão 5.5.3

As notas da versão a seguir incluem informações para a versão 5.5.3 do Amazon EMR. As alterações são referentes à versão 5.5.2.

Data da versão inicial: 29 de agosto de 2018

Alterações, melhorias e problemas resolvidos

- Esta versão aborda uma possível vulnerabilidade de segurança.

Versão 5.5.2

As notas da versão a seguir incluem informações para a versão 5.5.2 do Amazon EMR. As alterações são referentes à versão 5.5.1.

Data da versão inicial: 29 de março de 2018

Alterações, melhorias e problemas resolvidos

- Atualizado o kernel do Amazon Linux da AMI padrão do Amazon Linux para Amazon EMR para abordar possíveis vulnerabilidades.

Versão 5.5.1

As notas da versão a seguir incluem informações para a versão 5.5.1 do Amazon EMR. As alterações são referentes à versão 5.5.0 do Amazon EMR.

Data da versão inicial: 22 de janeiro de 2018

Alterações, melhorias e problemas resolvidos

- Atualizado o kernel do Amazon Linux da AMI padrão do Amazon Linux para Amazon EMR para abordar vulnerabilidades associadas à execução especulativa (CVE-2017-5715, CVE-2017-5753 e CVE-2017-5754). Para obter mais informações, consulte https://aws.amazon.com/security/security-bulletins/AWS-2018-013/.

Versão 5.5.0

As notas da versão a seguir incluem informações para a versão 5.5.0 do Amazon EMR. As alterações são referentes à versão 5.4.0 do Amazon EMR.

Data do release: 26 de abril de 2017

Atualizações

- Hue 3.12
- Presto 0.170
- Zeppelin 0.7.1

- ZooKeeper 3.4.10

Alterações, melhorias e problemas resolvidos

- Spark
  - Correção do Backported Spark Patch (SPARK-20115) DAGScheduler para recalcular todos os blocos de reprodução aleatória perdidos quando o serviço externo de reprodução aleatória não está disponível na versão 2.1.0 do Spark, incluída nesta versão.
- Flink
  - O Flink agora é compilado com o Scala 2.11. Se você usar a API e as bibliotecas do Scala, recomendamos que use o Scala 2.11 em seus projetos.
  - Tratado um problema em que os padrões `HADOOP_CONF_DIR` e `YARN_CONF_DIR` não estavam definidos corretamente, portanto havia falha no funcionamento de `start-scala-shell.sh`. Também foi adicionada a possibilidade de definir esses valores usando `env.hadoop.conf.dir` e `env.yarn.conf.dir` em `/etc/flink/conf/flink-conf.yaml` ou na classificação de configuração `flink-conf`.
  - Introduzido um novo comando específico do EMR, `flink-scala-shell`, como um wrapper para `start-scala-shell.sh`. Recomendamos o uso desse comando, em vez de `start-scala-shell`. O novo comando simplifica a execução. Por exemplo, `flink-scala-shell -n 2` inicia um shell Scala Flink com um paralelismo de tarefa de 2.
  - Introduzido um novo comando específico do EMR, `flink-yarn-session`, como um wrapper para `yarn-session.sh`. Recomendamos o uso desse comando, em vez de `yarn-session`. O novo comando simplifica a execução. Por exemplo, `flink-yarn-session -d -n 2` inicia uma sessão de longa execução do Flink em um estado desanexado com dois gerenciadores de tarefas.
  - Resolvido (FLINK-6125) commons httpclient não está mais esmaecido no Flink 1.2.
- Presto
  - Adicionado o suporte para autenticação do LDAP. Usar LDAP com o Presto no Amazon EMR requer que você habilite o acesso HTTPS para o coordenador do Presto (`http-server.https.enabled=true` em `config.properties`). Para obter detalhes da configuração, consulte Autenticação LDAP na documentação do Presto.
  - O suporte adicionado para `SHOW GRANTS`.
- AMI base do Linux do Amazon EMR

- As versões do Amazon EMR agora são baseadas no Amazon Linux 2017.03. Para obter mais informações, consulte Notas da versão da AMI do Amazon Linux 2017.03.
- Removido o Python 2.6 da imagem de base Linux do Amazon EMR. Python 2.7 e 3.4 estão instalados por padrão. Você pode instalar o Python 2.6 manualmente, se necessário.

Versão 5.4.0

As notas da versão a seguir incluem informações para a versão 5.4.0 do Amazon EMR. As alterações são referentes à versão 5.3.0 do Amazon EMR.

Data do release: 08 de março de 2017

Atualizações

As seguintes atualizações estão disponíveis nesta versão:

- Atualizado para Flink 1.2.0
- Atualizado para Hbase 1.3.0
- Atualizado para Phoenix 4.9.0

> **Note**
>
> Se você fez a atualização de uma versão mais antiga do Amazon EMR para a versão 5.4.0 ou posterior do Amazon EMR e usa indexação secundária, atualize os índices locais conforme descrito na documentação do Apache Phoenix. O Amazon EMR remove as configurações necessárias da classificação do `hbase-site`, mas os índices precisam ser preenchidos novamente. O sistema oferece suporte a atualizações de índices online e offline. As atualizações online são o padrão, o que significa que os índices são preenchidos novamente durante a inicialização de clientes do Phoenix versão 4.8.0 ou posterior. Para especificar as atualizações offline, defina a configuração do `phoenix.client.localIndexUpgrade` como falsa na classificação do `phoenix-site` e, em seguida, execute o SSH no nó principal para executar o `psql [zookeeper] -1`.

- Atualizado para Presto 0.166
- Atualizado para Zeppelin 0.7.0

Alterações e melhorias

As seguintes alterações foram feitas nas versões do Amazon EMR com o rótulo de versão emr-5.4.0:

- Adicionado o suporte para instâncias r4. Veja os tipos de EC2 instância da Amazon.

Versão 5.3.1

As notas da versão a seguir incluem informações para a versão 5.3.1 do Amazon EMR. As alterações são referentes à versão 5.3.0 do Amazon EMR.

Data da versão: 7 de fevereiro de 2017

Pequenas alterações nos patches do Zeppelin de backport e na atualização da AMI padrão para o Amazon EMR.

Versão 5.3.0

As notas da versão a seguir incluem informações para a versão 5.3.0 do Amazon EMR. As alterações são referentes à versão 5.2.1 do Amazon EMR.

Data do release: 26 de janeiro de 2017

Atualizações

As seguintes atualizações estão disponíveis nesta versão:

- Atualizado para Hive 2.1.1
- Atualizado para Hue 3.11.0
- Atualizado para Spark 2.1.0
- Atualizado para Oozie 4.3.0
- Atualizado para Flink 1.1.4

Alterações e melhorias

As seguintes alterações foram feitas nas versões do Amazon EMR com o rótulo de versão emr-5.3.0:

- Adicionado um patch para o Hue que permite usar a configuração `interpreters_shown_on_wheel` para definir o que intérpretes mostram primeiro na roda de seleção do bloco de anotações, independentemente de sua ordem no arquivo `hue.ini`.
- Adicionada a classificação de configuração `hive-parquet-logging`, que pode ser usada para configurar os valores no arquivo `parquet-logging.properties` do Hive.

Versão 5.2.2

As notas da versão a seguir incluem informações para a versão 5.2.2 do Amazon EMR. As alterações são referentes à versão 5.2.1 do Amazon EMR.

Data do release: 2 de maio de 2017

Problemas conhecidos das versões anteriores que foram resolvidos

- Backported SPARK-194459, que soluciona um problema em que a leitura de uma tabela ORC com colunas pode falhar. char/varchar

Versão 5.2.1

As notas da versão a seguir incluem informações para a versão 5.2.1 do Amazon EMR. As alterações são referentes à versão 5.2.0 do Amazon EMR.

Data do release: 29 de dezembro de 2016

Atualizações

As seguintes atualizações estão disponíveis nesta versão:

- Atualizado para Presto 0.157.1. Para obter mais informações, consulte Notas de versão do Presto na documentação do Presto.
- Zookeeper atualizado para 3.4.9. Para obter mais informações, consulte as notas de ZooKeeper lançamento na ZooKeeper documentação do Apache.

Alterações e melhorias

As seguintes alterações foram feitas nas versões do Amazon EMR com o rótulo de versão emr-5.2.1:

- Foi adicionado suporte para o tipo de instância Amazon EC2 m4.16xlarge no Amazon EMR versão 4.8.3 e posterior, excluindo 5.0.0, 5.0.3 e 5.2.0.
- As versões do Amazon EMR agora são baseadas no Amazon Linux 2016.09. Para obter mais informações, consulte https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/.
- A localização de caminhos de configuração do Flink e do YARN agora são definidas por padrão em `/etc/default/flink` para que você não precise definir as variáveis de ambiente `FLINK_CONF_DIR` e `HADOOP_CONF_DIR` ao executar os scripts de driver `flink` ou `yarn-session.sh` para iniciar trabalhos do Flink.
- Foi adicionado suporte para a FlinkKinesisConsumer aula.

Problemas conhecidos das versões anteriores que foram resolvidos

- Corrigido um problema no Hadoop em que o ReplicationMonitor thread podia ficar preso por um longo tempo devido a uma corrida entre a replicação e a exclusão do mesmo arquivo em um grande cluster.
- Corrigido um problema em que ControlledJob #toString falhava com uma exceção de ponteiro nulo (NPE) quando o status do trabalho não era atualizado com êxito.

Versão 5.2.0

As notas da versão a seguir incluem informações para a versão 5.2.0 do Amazon EMR. As alterações são referentes à versão 5.1.0 do Amazon EMR.

Data do release: 21 de novembro de 2016

Alterações e melhorias

As seguintes alterações e melhorias estão disponíveis nesta versão:

- Adicionado o modo de armazenamento Amazon S3 para. HBase
- Permite que você especifique uma localização do Amazon S3 para o HBase rootdir. Para obter mais informações, consulte HBaseno Amazon S3.

Atualizações

As seguintes atualizações estão disponíveis nesta versão:

- Atualizado para Spark 2.0.2

Problemas conhecidos das versões anteriores que foram resolvidos

- Corrigido um problema com /mnt ficar restrito a 2 TB em tipos de instância somente EBS.
- Corrigido um problema com o controlador de instância e com os logs do logpusher serem a saída para seus arquivos .out correspondentes, em vez de para seus arquivos normais .log configurados com log4j, que mudam de hora em hora. Os arquivos .out não mudam, portanto isso eventualmente encheria a partição /emr. Esse problema afeta somente tipos de instância de HVM (máquina virtual de hardware).

Versão 5.1.0

As notas da versão a seguir incluem informações para a versão 5.1.0 do Amazon EMR. As alterações são referentes à versão 5.0.0 do Amazon EMR.

Data do release: 03 de novembro de 2016

Alterações e melhorias

As seguintes alterações e melhorias estão disponíveis nesta versão:

- Adicionado o suporte para Flink 1.1.3.
- O Presto foi adicionado como uma opção na seção bloco de anotações do Hue.

Atualizações

As seguintes atualizações estão disponíveis nesta versão:

- Atualizado para 1.2.3 HBase
- Atualizado para Zeppelin 0.6.2

Problemas conhecidos das versões anteriores que foram resolvidos

- Corrigido um problema com consultas do Tez no Amazon S3 com arquivos ORC que não tinham uma performance tão boa quanto em versões 4.x anteriores do Amazon EMR.

Versão 5.0.3

As notas da versão a seguir incluem informações para a versão 5.0.3 do Amazon EMR. As alterações são referentes à versão 5.0.0 do Amazon EMR.

Data do release: 24 de outubro de 2016

Atualizações

As seguintes atualizações estão disponíveis nesta versão:

- Atualizado para Hadoop 2.7.3
- Atualizado para Presto 0.152.3, que inclui o suporte para a interface da web do Presto. Você pode acessar a interface da web do Presto no coordenador do Presto usando a porta 8889. Para obter mais informações sobre a interface da Web do Presto, consulte Interface da Web na documentação do Presto.
- Atualizado para Spark 2.0.1
- As versões do Amazon EMR agora são baseadas no Amazon Linux 2016.09. Para obter mais informações, consulte https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/.

Versão 5.0.0

Data do release: 27 de julho de 2016

Atualizações

As seguintes atualizações estão disponíveis nesta versão:

- Atualizado para Hive 2.1
- Atualizado para Presto 0.150
- Atualizado para Spark 2.0
- Atualizado para Hue 3.10.0
- Atualizado para Pig 0.16.0
- Atualizado para Tez 0.8.4
- Atualizado para Zeppelin 0.6.1

Alterações e melhorias

As seguintes alterações foram feitas nas versões do Amazon EMR com o rótulo de versão emr-5.0.0 ou posterior:

- O Amazon EMR é compatível com as versões de código aberto mais recentes do Hive (versão 2.1) e do Pig (versão 0.16.0). Se, anteriormente, você tiver usado o Hive ou Pig no Amazon EMR, isso poderá afetar alguns casos de uso. Para obter mais informações, consulte Hive e Pig.
- O mecanismo de execução padrão para o Hive e o Pig agora é o Tez. Para alterar isso, você deve editar os valores apropriados nas classificações de configuração `hive-site` e `pig-properties`, respectivamente.
- Um recurso de etapa aprimorada de depuração foi adicionado, o que permite que você veja a causa raiz de falhas de etapa se o serviço puder determinar a causa. Para obter mais informações, consulte Depuração de etapa aprimorada no Guia de gerenciamento do Amazon EMR.
- Os aplicativos que, anteriormente, terminavam com "-Sandbox" não têm mais esse sufixo. Isso pode inutilizar sua automação, por exemplo, se você estiver usando scripts para iniciar clusters com esses aplicativos. A tabela a seguir mostra nomes de aplicações na versão 4.7.2 do Amazon EMR em relação à versão 5.0.0 do Amazon EMR.

  Alterações dos nomes de aplicativos

Amazon EMR 4.7.2	Amazon EMR 5.0.0
Oozie-Sandbox	Oozie
Presto-Sandbox	Presto
Sqoop-Sandbox	Sqoop
Zeppelin-Sandbox	Zeppelin
ZooKeeper-Caixa de areia	ZooKeeper

- O Spark agora está compilado para Scala 2.11.
- Java 8 é agora o JVM padrão. Todas as aplicações são executadas usando o runtime do Java 8. Não há alterações em qualquer destino de código de bytes da aplicação. A maioria dos aplicativos continuam a usar o Java 7 como destino.
- O Zeppelin agora inclui recursos de autenticação. Para obter mais informações, consulte Zeppelin.

- Adicionado o suporte para configurações de segurança, que permitem criar e aplicar opções de criptografia com mais facilidade. Para obter mais informações, consulte Criptografia de dados.

Versão 4.9.5

As notas da versão a seguir incluem informações para a versão 4.9.5 do Amazon EMR. As alterações são referentes à versão 4.9.4.

Data da versão inicial: 29 de agosto de 2018

Alterações, melhorias e problemas resolvidos

- HBase
  - Esta versão aborda uma possível vulnerabilidade de segurança.

Versão 4.9.4

As notas da versão a seguir incluem informações para a versão 4.9.4 do Amazon EMR. As alterações são referentes à versão 4.9.3.

Data da versão inicial: 29 de março de 2018

Alterações, melhorias e problemas resolvidos

- Atualizado o kernel do Amazon Linux da AMI padrão do Amazon Linux para Amazon EMR para abordar possíveis vulnerabilidades.

Versão 4.9.3

As notas da versão a seguir incluem informações para a versão 4.9.3 do Amazon EMR. As alterações são referentes à versão 4.9.2 do Amazon EMR.

Data da versão inicial: 22 de janeiro de 2018

Alterações, melhorias e problemas resolvidos

- Atualizado o kernel do Amazon Linux da AMI padrão do Amazon Linux para Amazon EMR para abordar vulnerabilidades associadas à execução especulativa (CVE-2017-5715, CVE-2017-5753 e

CVE-2017-5754). Para obter mais informações, consulte https://aws.amazon.com/security/security-bulletins/AWS-2018-013/.

Versão 4.9.2

As notas da versão a seguir incluem informações para a versão 4.9.2 do Amazon EMR. As alterações são referentes à versão 4.9.1 do Amazon EMR.

Data do release: 13 de julho de 2017

Pequenas alterações, correções de erros e melhorias foram feitas nesta versão.

Versão 4.9.1

As notas da versão a seguir incluem informações para a versão 4.9.1 do Amazon EMR. As alterações são referentes à versão 4.8.4 do Amazon EMR.

Data do release: 10 de abril de 2017

Problemas conhecidos das versões anteriores que foram resolvidos

- Envio para backport do HIVE-9976 e do HIVE-10106
- Corrigido um problema no YARN em que um grande número de nós (mais de 2.000) e de contêineres (mais de 5.000) causavam um erro de memória insuficiente, por exemplo: `"Exception in thread 'main' java.lang.OutOfMemoryError"`.

Alterações e melhorias

As seguintes alterações foram feitas nas versões do Amazon EMR com o rótulo de versão emr-4.9.1:

- As versões do Amazon EMR agora são baseadas no Amazon Linux 2017.03. Para obter mais informações, consulte https://aws.amazon.com/amazon-linux-ami/2017.03-release-notes/.
- Removido o Python 2.6 da imagem de base Linux do Amazon EMR. Você pode instalar o Python 2.6 manualmente, se necessário.

Versão 4.8.4

As notas da versão a seguir incluem informações para a versão 4.8.4 do Amazon EMR. As alterações são referentes à versão 4.8.3 do Amazon EMR.

Data do release: 7 de fevereiro de 2017

Pequenas alterações, correções de erros e melhorias foram feitas nesta versão.

Versão 4.8.3

As notas da versão a seguir incluem informações para a versão 4.8.3 do Amazon EMR. As alterações são referentes à versão 4.8.2 do Amazon EMR.

Data do release: 29 de dezembro de 2016

Atualizações

As seguintes atualizações estão disponíveis nesta versão:

- Atualizado para Presto 0.157.1. Para obter mais informações, consulte Notas de versão do Presto na documentação do Presto.
- Atualizado para Spark 1.6.3. Para obter mais informações, consulte Spark release notes na documentação do Apache Spark.
- Atualizado para ZooKeeper 3.4.9. Para obter mais informações, consulte as notas de ZooKeeper lançamento na ZooKeeper documentação do Apache.

Alterações e melhorias

As seguintes alterações foram feitas nas versões do Amazon EMR com o rótulo de versão emr-4.8.3:

- Foi adicionado suporte para o tipo de instância Amazon EC2 m4.16xlarge no Amazon EMR versão 4.8.3 e posterior, excluindo 5.0.0, 5.0.3 e 5.2.0.
- As versões do Amazon EMR agora são baseadas no Amazon Linux 2016.09. Para obter mais informações, consulte https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/.

Problemas conhecidos das versões anteriores que foram resolvidos

- Corrigido um problema no Hadoop em que o ReplicationMonitor thread podia ficar preso por um longo tempo devido a uma corrida entre a replicação e a exclusão do mesmo arquivo em um grande cluster.
- Corrigido um problema em que ControlledJob #toString falhava com uma exceção de ponteiro nulo (NPE) quando o status do trabalho não era atualizado com êxito.

Versão 4.8.2

As notas da versão a seguir incluem informações para a versão 4.8.2 do Amazon EMR. As alterações são referentes à versão 4.8.0 do Amazon EMR.

Data do release: 24 de outubro de 2016

Atualizações

As seguintes atualizações estão disponíveis nesta versão:

- Atualizado para Hadoop 2.7.3
- Atualizado para Presto 0.152.3, que inclui o suporte para a interface da web do Presto. Você pode acessar a interface da web do Presto no coordenador do Presto usando a porta 8889. Para obter mais informações sobre a interface da Web do Presto, consulte Interface da Web na documentação do Presto.
- As versões do Amazon EMR agora são baseadas no Amazon Linux 2016.09. Para obter mais informações, consulte https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/.

Versão 4.8.0

Data do release: 7 de setembro de 2016

Atualizações

As seguintes atualizações estão disponíveis nesta versão:

- Atualizado para 1.2.2 HBase
- Atualizado para Presto-Sandbox 0.151
- Atualizado para Tez 0.8.4
- Atualizado para Zeppelin-Sandbox 0.6.1

Alterações e melhorias

As seguintes alterações foram feitas nas versões do Amazon EMR com o rótulo de versão emr-4.8.0:

- Corrigido um problema no YARN em que eles ApplicationMaster tentavam limpar contêineres que não existem mais porque suas instâncias foram encerradas.

- Corrigido o URL do hive-server2 para ações do Hive2 em exemplos do Oozie.
- Adicionado o suporte para catálogos Presto adicionais.
- Patches enviados para backport: HIVE-8948, HIVE-12679, HIVE-13405, PHOENIX-3116, HADOOP-12689
- Adicionado o suporte para configurações de segurança, que permitem criar e aplicar opções de criptografia com mais facilidade. Para obter mais informações, consulte Criptografia de dados.

Versão 4.7.2

As notas da versão a seguir incluem informações para a versão 4.7.2 do Amazon EMR.

Data do release: 15 de julho de 2016

Atributos

Os seguintes recursos estão disponíveis nesta versão:

- Atualizado para Mahout 0.12.2
- Atualizado para Presto 0.148
- Atualizado para Spark 1.6.2
- Agora você pode criar um AWSCredentials provedor para uso com o EMRFS usando um URI como parâmetro. Para obter mais informações, consulte Criar um AWSCredentials provedor para o EMRFS.
- O EMRFS agora permite que os usuários configurem um endpoint personalizado do Dynamo para seus metadados de visualização consistente usando a propriedade `fs.s3.consistent.dynamodb.endpoint` em `emrfs-site.xml`.
- Adicionado um script em `/usr/bin` chamado `spark-example`, que encapsula `/usr/lib/spark/spark/bin/run-example` para que você possa executar exemplos diretamente. Por exemplo, para executar o SparkPi exemplo que vem com a distribuição do Spark, você pode executar a `spark-example SparkPi 100` partir da linha de comando ou usando `command-runner.jar` como uma etapa na API.

Problemas conhecidos das versões anteriores que foram resolvidos

- Corrigido um problema em que o Oozie não tinha o `spark-assembly.jar` no local correto quando o Spark também estava instalado, o que resultava em falha para iniciar aplicativos do Spark com o Oozie.
- Corrigido um problema com o registro baseado em Log4j do Spark em contêineres do YARN.

Versão 4.7.1

Data do release: 10 de junho de 2016

Problemas conhecidos das versões anteriores que foram resolvidos

- Corrigido um problema que estendia o tempo de inicialização de clusters iniciados em uma VPC com sub-redes privadas. O bug só afetava os clusters iniciados com a versão 4.7.0. do Amazon EMR.
- Corrigido um problema que tratava incorretamente a listagem de arquivos no Amazon EMR para os clusters iniciados com a versão 4.7.0 do Amazon EMR.

Versão 4.7.0

> **Important**
>
> A versão 4.7.0 do Amazon EMR foi descontinuada. Use a versão 4.7.1 ou posterior do Amazon EMR.

Data do release: 2 de junho de 2016

Atributos

Os seguintes recursos estão disponíveis nesta versão:

- Adicionado o Apache Phoenix 4.7.0
- Adicionado o Apache Tez 0.8.3
- Atualizado para 1.2.1 HBase

- Atualizado para Mahout 0.12.0
- Atualizado para Presto 0.147
- Atualizou o AWS SDK para Java para 1.10.75
- O sinalizador final foi removido da propriedade `mapreduce.cluster.local.dir` em `mapred-site.xml` para permitir que os usuários executem o Pig no modo local.

Drivers JDBC do Amazon Redshift disponíveis no cluster

Os drivers JDBC do Amazon Redshift agora estão incluídos em `/usr/share/aws/redshift/jdbc`. `/usr/share/aws/redshift/jdbc/RedshiftJDBC41.jar` é o driver do Amazon Redshift compatível com JDBC 4.1 e `/usr/share/aws/redshift/jdbc/RedshiftJDBC4.jar` é o driver do Amazon Redshift compatível com JDBC 4.0. Para obter mais informações, consulte Configurar uma conexão JDBC no Guia de gerenciamento do Amazon Redshift.

Java 8

Com exceção do Presto, o OpenJDK 1.7 é o JDK padrão usado para todos os aplicativos. No entanto, o OpenJDK 1.7 e o 1.8 estão instalados. Para obter mais informações sobre como configurar `JAVA_HOME` para aplicações, consulte Configurar aplicações para usar Java 8.

Problemas conhecidos das versões anteriores que foram resolvidos

- Corrigido um problema do kernel que afetava significativamente a performance dos volumes do EBS no disco rígido com throughput otimizado (st1) para o Amazon EMR no emr-4.6.0.
- Corrigido um problema em que havia uma falha do cluster se qualquer zona de criptografia do HDFS fosse especificada sem escolher o Hadoop como um aplicativo.
- Alterada a política de gravação padrão do HDFS de `RoundRobin` para `AvailableSpaceVolumeChoosingPolicy`. Alguns volumes não foram utilizados adequadamente com a RoundRobin configuração, o que resultou em falhas nos nós principais e em um HDFS não confiável.
- Corrigido um problema com a CLI do EMRFS, que causava uma exceção ao criar a tabela de metadados do DynamoDB padrão para visualizações consistentes.
- Corrigido um problema de deadlock no EMRFS que potencialmente ocorria durante a renomeação do multipart e operações de cópia.
- Corrigido um problema com o EMRFS que fazia com que o CopyPart tamanho padrão fosse 5 MB. O padrão agora está definido corretamente como 128 MB.

- Corrigido um problema com a configuração de inicialização do Zeppelin que potencialmente impedia a interrupção do serviço.
- Corrigido um problema com o Spark e o Zeppelin que impedia o uso do esquema de URI `s3a://` porque `/usr/lib/hadoop/hadoop-aws.jar` não era carregado adequadamente em seus respectivos classpath.
- HUE-2484 enviado para backport.
- Reportou um commit do Hue 3.9.0 (não existe JIRA) para corrigir um problema com a amostra do navegador. HBase
- HIVE-9073 enviado para backport.

Versão 4.6.0

Data do release: 21 de abril de 2016

Atributos

Os seguintes recursos estão disponíveis nesta versão:

- Adicionado HBase 1.2.0
- Adicionado o ZooKeeper-Sandbox 3.4.8
- Atualizado para Presto-Sandbox 0.143
- As versões do Amazon EMR agora são baseadas no Amazon Linux 2016.03.0. Para obter mais informações, consulte https://aws.amazon.com/amazon-linux-ami/2016.03-release-notes/.

Problema afetando os tipos de volumes do EBS no disco rígido com throughput otimizado (st1)

Um problema no kernel do Linux versões 4.2 e superior afeta significativamente o desempenho nos volumes do EBS no disco rígido com throughput otimizado (st1) para o EMR. Esta versão (emr-4.6.0) usa uma versão do kernel 4.4.5 e, portanto, é afetada. Por isso, recomendamos não usar o emr-4.6.0 se você deseja usar volumes do EBS no st1. Você pode usar emr-4.5.0 ou versões anteriores do Amazon EMR com o st1 sem impacto algum. Além disso, fornecemos a correção com futuras versões.

Padrões do Python

O Python 3.4 agora está instalado por padrão, mas o Python 2.7 permanece como o sistema padrão. Você pode configurar o Python 3.4 como o padrão do sistema usando uma ação de bootstrap; você pode usar a API de configuração para definir a exportação de PYSPARK_PYTHON na classificação `spark-env` para afetar a versão `/usr/bin/python3.4` do Python usada por. PySpark

Java 8

Com exceção do Presto, o OpenJDK 1.7 é o JDK padrão usado para todos os aplicativos. No entanto, o OpenJDK 1.7 e o 1.8 estão instalados. Para obter mais informações sobre como configurar `JAVA_HOME` para aplicações, consulte Configurar aplicações para usar Java 8.

Problemas conhecidos das versões anteriores que foram resolvidos

- Corrigido um problema em que, às vezes, ocorria uma falha aleatória no provisionamento de aplicativos devido a uma senha gerada.
- Anteriormente, `mysqld` estava instalado em todos os nós. Agora, ele só está instalado na instância principal e somente se o aplicativo escolhido incluir `mysql-server` como componente. Atualmente, os seguintes aplicativos incluem o `mysql-server` componente: Hive HCatalog, Hue, Presto-Sandbox e Sqoop-Sandbox.
- Alterado o `yarn.scheduler.maximum-allocation-vcores` para 80 do padrão de 32, que corrige um problema apresentado na emr-4.4.0 que ocorre, principalmente, com o Spark enquanto usa a opção `maximizeResourceAllocation` em um cluster cujo tipo de instância core é um dos grandes tipos de instância que têm os vcores do YARN definidos como maior do que 32, ou seja, c4.8xlarge, cc2.8xlarge, hs1.8xlarge, i2.8xlarge, m2.4xlarge, r3.8xlarge, d2.8xlarge ou m4.10xlarge foram afetados por esse problema.
- s3-dist-cp agora usa o EMRFS para todas as indicações do Amazon S3 e não mais estágios para um diretório HDFS temporário.
- Corrigido um problema com o tratamento de exceções para os multipart uploads de criptografia no lado do cliente.
- Adicionada uma opção para permitir que os usuários alterem a classe de armazenamento do Amazon S3. Por padrão, essa configuração é `STANDARD`. A configuração da classificação de configuração `emrfs-site` é `fs.s3.storageClass` e os valores possíveis são `STANDARD`, `STANDARD_IA`e `REDUCED_REDUNDANCY`. Para obter mais informações sobre classes de armazenamento, consulte Classes de armazenamento no Guia do usuário do Amazon Simple Storage Service.

Versão 4.5.0

Data do release: 4 de abril de 2016

Atributos

Os seguintes recursos estão disponíveis nesta versão:

- Atualizado para Spark 1.6.1
- Atualizado para Hadoop 2.7.2
- Atualizado para Presto 0.140
- Foi adicionado AWS KMS suporte para criptografia do lado do servidor Amazon S3.

Problemas conhecidos das versões anteriores que foram resolvidos

- Corrigido um problema em que os servidores MySQL e Apache não iniciavam depois que um nó fosse reinicializado.
- Corrigido um problema em que IMPORT não funcionava corretamente com tabelas não particionadas armazenadas no Amazon S3
- Corrigido um problema em que o Presto exigia que o diretório de preparo fosse `/mnt/tmp` em vez de `/tmp` ao gravar em tabelas do Hive.

Versão 4.4.0

Data do release: 14 de março de 2016

Atributos

Os seguintes recursos estão disponíveis nesta versão:

- Adicionado HCatalog 1.0.0
- Adicionado o Sqoop-Sandbox 1.4.6
- Atualizado para Presto 0.136
- Atualizado para Zeppelin 0.5.6
- Atualizado para Mahout 0.11.1
- Habilitada a `dynamicResourceAllocation` por padrão.

- Adicionada uma tabela de todas as classificações de configuração para a versão. Para obter mais informações, consulte a tabela de classificações de configuração em Configurar aplicações.

Problemas conhecidos das versões anteriores que foram resolvidos

- Corrigido um problema em que a `maximizeResourceAllocation` configuração não reservava memória suficiente para os ApplicationMaster daemons do YARN.
- Corrigido um problema encontrado com um DNS personalizado. Se alguma entrada em `resolve.conf` preceder as entradas personalizadas fornecidas, as entradas personalizadas não serão resolvíveis. Esse comportamento foi afetado por clusters em uma VPC em que o servidor de nomes padrão da VPC está inserido como a entrada superior em `resolve.conf`.
- Corrigido um problema em que o Python padrão mudou para a versão 2.7 e boto não estava instalado para essa versão.
- Corrigido um problema em que contêineres do YARN e aplicativos do Spark geravam um arquivo (rrd) exclusivo do banco de dados Ganglia round robin, que resultava no enchimento do primeiro disco anexado à instância. Devido a essa correção, as métricas no nível do contêiner do YARN e as métricas no nível do aplicativo do Spark foram desativadas.
- Corrigido um problema no log pusher em que ele excluía todas as pastas de log vazias. O efeito era que a CLI do Hive não podia criar logs porque o log pusher estava removendo a pasta `user` vazia em `/var/log/hive`.
- Corrigido um problema com as importações do Hive, que afetava o particionamento e resultava em um erro durante a importação.
- Corrigido um problema em que EMRFS e s3-dist-cp não lidavam adequadamente com nomes de bucket que continham pontos.
- Alterado um comportamento no EMRFS, de modo que nos buckets habilitados para versionamento, o arquivo marcador `_$folder$` não é criado continuamente, o que pode contribuir para um melhor desempenho dos buckets habilitados para versionamento.
- Alterado o comportamento no EMRFS de forma que ele não use os arquivos de instrução, exceto para casos em que a criptografia no lado do cliente esteja habilitada. Se desejar excluir os arquivos de instrução enquanto usa a criptografia no lado do cliente, você poderá definir a propriedade emrfs-site.xml `fs.s3.cse.cryptoStorageMode.deleteInstructionFiles.enabled`, como verdadeiro.
- Alterada a agregação de logs do YARN para reter os logs no destino de agregação por dois dias. O destino padrão é o armazenamento HDFS do cluster. Se desejar mudar essa duração, altere o

valor de `yarn.log-aggregation.retain-seconds` usando a classificação de configuração `yarn-site` quando criar seu cluster. Como sempre, você pode salvar os logs de aplicações no Amazon S3 usando o parâmetro `log-uri` quando criar o cluster.

Patches aplicados

Os seguintes patches de projetos de código aberto foram incluídos nesta versão:

- HIVE-9655
- HIVE-9183
- HADOOP-12810

Versão 4.3.0

Data do release: 19 de janeiro de 2016

Atributos

Os seguintes recursos estão disponíveis nesta versão:

- Atualizado para Hadoop 2.7.1
- Atualizado para Spark 1.6.0
- Ganglia atualizado para 3.7.2
- Presto atualizado para 0.130

O Amazon EMR fez algumas alterações na configuração `spark.dynamicAllocation.enabled` quando ela estava definida como verdadeiro. Ela é definida como falso por padrão.
Quando definida como verdadeiro, isso afeta os padrões definidos pela configuração `maximizeResourceAllocation`:

- Se `spark.dynamicAllocation.enabled` estiver definida como true, `spark.executor.instances` não será definida por `maximizeResourceAllocation`.
- A configuração `spark.driver.memory` agora é definida com base nos tipos de instância no cluster de maneira semelhante a como `spark.executors.memory` é definida. No entanto, como o aplicativo de driver do Spark pode ser executado na instância principal ou em uma das instâncias core (por exemplo, nos modos de cluster e de cliente do YARN, respectivamente), a configuração

`spark.driver.memory` é definida com base no tipo de instância do menor tipo de instância entre esses dois grupos de instâncias.

- A configuração `spark.default.parallelism` agora é definida em duas vezes mais núcleos de CPU disponíveis para contêineres do YARN. Em versões anteriores, era a metade desse valor.
- Os cálculos para a sobrecarga de memória reservada para os processos do Spark YARN foram ajustados para serem mais precisos, resultando em um pequeno aumento na quantidade total de memória disponível para o Spark (ou seja, `spark.executor.memory`).

Problemas conhecidos das versões anteriores que foram resolvidos

- A agregação de logs do YARN agora está ativada por padrão.
- Corrigido um problema em que os logs não eram enviados para um bucket de logs do Amazon S3 do cluster quando a agregação de logs do YARN estava habilitada.
- Os contêineres do YARN agora têm um novo mínimo de 32 em todos os tipos de nós.
- Corrigido um problema com o Ganglia que causava excesso de disco I/O no nó principal em grandes clusters.
- Corrigido um problema que impedia que os logs de aplicações fossem enviados para o Amazon S3 quando um cluster estivesse sendo encerrado.
- Corrigido um problema na CLI do EMRFS que fazia com que houvesse falha em determinados comandos.
- Corrigido um problema com o Zeppelin que impedia que dependências fossem carregadas no subjacente. SparkContext
- Corrigido um problema resultante da emissão de um redimensionamento para tentar adicionar instâncias.
- Corrigido um problema no Hive, em que CREATE TABLE AS SELECT fazia chamadas de lista excessivas para o Amazon S3.
- Corrigido um problema em que clusters grandes não provisionavam corretamente quando o Hue, o Oozie e o Ganglia estivessem instalados.
- Corrigido um problema no s3-dist-cp em que retornava um código de saída zero, mesmo se houvesse falha com um erro.

Patches aplicados

Os seguintes patches de projetos de código aberto foram incluídos nesta versão:

- OOZIE-2402
- HIVE-12502
- HIVE-10631
- HIVE-12213
- HIVE-10559
- HIVE-12715
- HIVE-10685

Versão 4.2.0

Data do release: 18 de novembro de 2015

Atributos

Os seguintes recursos estão disponíveis nesta versão:

- Adicionado o suporte ao Ganglia
- Atualizado para Spark 1.5.2
- Atualizado para Presto 0.125
- Oozie atualizado para 4.2.0
- Zeppelin atualizado para 0.5.5
- Atualizou o AWS SDK para Java para 1.10.27

Problemas conhecidos das versões anteriores que foram resolvidos

- Corrigido um problema com a CLI do EMRFS em que ela não usava o nome da tabela de metadados padrão.
- Corrigido um problema encontrado ao serem usadas tabelas com ORC no Amazon S3.
- Corrigido um problema encontrado com uma divergência de versão do Python na configuração do Spark.
- Corrigido um problema quando havia falha de um status de nó YARN em relatar problemas de DNS para clusters em uma VPC.
- Corrigido um problema encontrado quando o YARN desativava nós, resultando em aplicativos suspensos ou na incapacidade de agendar novos aplicativos.

- Corrigido um problema encontrado quando os clusters eram encerrados com o status TIMED_OUT_STARTING.
- Corrigido um problema encontrado quando dependência Scala do EMRFS Scala era incluída em outras compilações. A dependência Scala foi removida.

Configurar aplicações

Para substituir as configurações padrão de uma aplicação, você pode fornecer um objeto de configuração. Você pode usar uma sintaxe abreviada para fornecer a configuração ou referenciar o objeto de configuração em um arquivo JSON. Os objetos de configuração consistem em uma classificação, propriedades e configurações opcionais aninhadas As propriedades correspondem às configurações da aplicação que você deseja alterar. Você pode especificar várias classificações para diversas aplicações em um único objeto JSON.

> **Warning**
>
> As operações da API Describe and List do Amazon EMR emitem configurações personalizadas e configuráveis, que são usadas como parte dos fluxos de trabalho do Amazon EMR, em texto se formatação. Para fornecer informações sigilosas, como senhas, nessas configurações, consulte Store sensitive configuration data in AWS Secrets Manager.

As classificações de configuração disponíveis variam de acordo com a versão do Amazon EMR. Para obter uma lista das classificações de configuração compatíveis com determinada versão, consulte a página da versão em Sobre as versões do Amazon EMR.

Veja a seguir um exemplo de arquivo JSON para uma lista de configurações.

```
[
  {
    "Classification": "core-site",
    "Properties": {
      "hadoop.security.groups.cache.secs": "250"
    }
  },
  {
    "Classification": "mapred-site",
    "Properties": {
      "mapred.tasktracker.map.tasks.maximum": "2",
      "mapreduce.map.sort.spill.percent": "0.90",
      "mapreduce.tasktracker.reduce.tasks.maximum": "5"
    }
  }
]
```

Uma classificação de configuração geralmente é mapeada para um arquivo de configuração específico do aplicativo. Por exemplo, a classificação `hive-site` é mapeada para as configurações no arquivo de configuração `hive-site.xml` para o Hive. Uma exceção para isso é a ação de bootstrap `configure-daemons`, que já não tem mais suporte, que é usada para definir parâmetros de ambiente, como `--namenode-heap-size`. Opções como esta são incluídas nas classificações `hadoop-env` e `yarn-env` com suas próprias classificações de exportação aninhadas. Se qualquer classificação terminar com `env`, use a subclassificação de exportação.

Outra exceção é `s3get`, que é usado para colocar um objeto `EncryptionMaterialsProvider` cliente em cada nó de um cluster para uso na criptografia do lado do cliente. Uma opção foi adicionada à classificação `emrfs-site` para essa finalidade.

Veja a seguir um exemplo da classificação `hadoop-env`.

```
[
  {
    "Classification": "hadoop-env",
    "Properties": {

    },
    "Configurations": [
      {
        "Classification": "export",
        "Properties": {
          "HADOOP_DATANODE_HEAPSIZE": "2048",
          "HADOOP_NAMENODE_OPTS": "-XX:GCTimeRatio=19"
        },
        "Configurations": [

        ]
      }
    ]
  }
]
```

Veja a seguir um exemplo da classificação yarn-env.

```
[
  {
    "Classification": "yarn-env",
    "Properties": {
```

```
    },
    "Configurations": [
      {
        "Classification": "export",
        "Properties": {
          "YARN_RESOURCEMANAGER_OPTS": "-Xdebug -Xrunjdwp:transport=dt_socket"
        },
        "Configurations": [

        ]
      }
    ]
  }
]
```

As configurações a seguir não pertencem a um arquivo de configuração, mas são usadas pelo Amazon EMR para configurar potencialmente várias configurações em seu nome.

Configurações selecionadas pelo Amazon EMR

Aplicação	Classificação do rótulo da versão	Propriedades válidas	Quando usar
Spark	`spark`	`maximizeResourceAllocation`	Configure executores para utilizar o máximo de recursos de cada nó.

Tópicos

- Configurar aplicações ao criar um cluster
- Reconfigurar um grupo de instâncias de um cluster em execução
- Armazenar dados de configuração sigilosos no AWS Secrets Manager
- Configurar aplicações para usar uma Máquina Virtual Java específica

Configurar aplicações ao criar um cluster

Ao criar um cluster, você pode substituir as configurações padrão das aplicações usando o console do Amazon EMR, a AWS Command Line Interface (AWS CLI) ou o SDK. AWS

Para substituir a configuração padrão de uma aplicação, especifique valores personalizados em uma classificação de configuração. Uma classificação de configuração corresponde a um arquivo XML de configuração de uma aplicação, como `hive-site.xml`.

As classificações de configuração variam de acordo com a versão do Amazon EMR. Para obter uma lista de classificações de configuração que estão disponíveis em uma versão específica, consulte a página de detalhes da versão. Por exemplo, o Amazon EMR versão 6.4.0.

Fornecer uma configuração no console ao criar um cluster

Para fornecer uma configuração, navegue até a página Criar cluster e expanda Configurações do software. Você poderá inserir a configuração diretamente usando JSON ou uma sintaxe abreviada demonstrada no texto sombreado no console. Caso contrário, você pode fornecer um URI do Amazon S3 para um arquivo com um objeto JSON `Configurations`.

Para fornecer uma configuração para um grupo de instâncias, escolha um cluster na lista de clusters e selecione a guia Configurações. Na tabela de Configurações do grupo de instâncias, escolha o grupo de instâncias a ser editado e, em seguida, Reconfigurar.

Fornecer uma configuração usando a AWS CLI ao criar um cluster

Você pode fornecer uma configuração para create-cluster fornecendo um caminho para um arquivo JSON armazenado localmente ou no Amazon S3. O exemplo a seguir pressupõe que você esteja usando perfis padrão para o Amazon EMR e que esses perfis foram criados. Se você precisar criar as funções, execute `aws emr create-default-roles` primeiro.

Se a configuração estiver em seu diretório local, você poderá usar o exemplo de comando a seguir.

```
aws emr create-cluster --use-default-roles --release-label emr-7.9.0 --applications
 Name=Hive \
--instance-type m5.xlarge --instance-count 3 --configurations file://./
configurations.json
```

Se a configuração estiver em um caminho do Amazon S3, você precisará configurar a seguinte solução alternativa antes de passar o caminho do Amazon S3 para o comando `create-cluster`.

```
#!/bin/sh
# Assume the ConfigurationS3Path is not public, and its present in the same AWS account
 as the EMR cluster
ConfigurationS3Path="s3://amzn-s3-demo-bucket/config.json"
# Get a presigned HTTP URL for the s3Path
```

```
ConfigurationURL=`aws s3 presign $ConfigurationS3Path --expires-in 300`
# Fetch the presigned URL, and minify the JSON so that it spans only a single line
Configurations=`curl $ConfigurationURL | jq -c .`
aws emr create-cluster --use-default-roles --release-label emr-5.34.0 --instance-type
 m5.xlarge --instance-count 2 --applications Name=Hadoop Name=Spark --configurations
 $Configurations
```

Fornecer uma configuração usando o SDK do Java ao criar um cluster

O trecho de programa a seguir mostra como fornecer uma configuração usando o AWS SDK para Java.

```
Application hive = new Application().withName("Hive");

Map<String,String> hiveProperties = new HashMap<String,String>();
 hiveProperties.put("hive.join.emit.interval","1000");
 hiveProperties.put("hive.merge.mapfiles","true");

Configuration myHiveConfig = new Configuration()
 .withClassification("hive-site")
 .withProperties(hiveProperties);

RunJobFlowRequest request = new RunJobFlowRequest()
 .withName("Create cluster with ReleaseLabel")
 .withReleaseLabel("emr-5.20.0")
 .withApplications(hive)
 .withConfigurations(myHiveConfig)
 .withServiceRole("EMR_DefaultRole")
 .withJobFlowRole("EMR_EC2_DefaultRole")
 .withInstances(new JobFlowInstancesConfig()
  .withEc2KeyName("myEc2Key")
  .withInstanceCount(3)
  .withKeepJobFlowAliveWhenNoSteps(true)
  .withMasterInstanceType("m4.large")
  .withSlaveInstanceType("m4.large")
 );
```

Reconfigurar um grupo de instâncias de um cluster em execução

Com o Amazon EMR versão 5.21.0 e posteriores, você pode reconfigurar as aplicações do cluster e especificar classificações de configuração adicionais para cada grupo de instâncias em um cluster

em execução. Para fazer isso, você pode usar o console do Amazon EMR, o AWS Command Line Interface (AWS CLI) ou o AWS SDK.

Quando você atualiza uma configuração de aplicação para um grupo de instâncias no novo console do Amazon EMR, o console tenta mesclar a nova configuração com a configuração atual para criar uma configuração nova e ativa. No caso incomum em que o Amazon EMR não conseguir mesclar a configuração, o console alertará você.

Depois que você enviar uma solicitação de reconfiguração para um grupo de instâncias, o Amazon EMR atribuirá um número de versão à nova especificação de configuração. É possível rastrear o número da versão de configuração ou o estado de um grupo de instâncias visualizando os CloudWatch eventos. Para obter mais informações, consulte Monitorar CloudWatch eventos.

> **Note**
>
> Só é possível substituir, e não excluir, as configurações do cluster que foram especificadas durante a criação do cluster. Se há diferenças entre a configuração atual e o arquivo que você forneceu, o Amazon EMR redefine as configurações modificadas manualmente, como as que você modificou enquanto estava conectado ao cluster usando SSH, para os padrões do cluster do grupo de instâncias especificado.

Considerações ao reconfigurar um grupo de instâncias

Ações de reconfiguração

Quando você envia uma solicitação de reconfiguração usando o console do Amazon EMR, AWS Command Line Interface a AWS CLI() ou o SDK, AWS o Amazon EMR verifica o arquivo de configuração existente no cluster. Se há diferenças entre a configuração atual e o arquivo que você forneceu, o Amazon EMR inicia ações de reconfiguração, reinicia algumas aplicações e redefine as configurações modificadas manualmente, como as que você modificou enquanto estava conectado ao cluster usando SSH, para os padrões do cluster do grupo de instâncias especificado.

> **Note**
>
> O Amazon EMR executa algumas ações padrão durante cada reconfiguração de grupo de instâncias. Essas ações padrão podem entrar em conflito com as personalizações do cluster feitas por você e resultar em falhas de reconfiguração. Para obter mais

informações sobre como solucionar falhas de reconfiguração, consulte Solucionar problemas de reconfiguração do grupo de instâncias.

O Amazon EMR também inicia ações de reconfiguração para as classificações de configuração que você especifica na solicitação. Para obter uma lista completa dessas ações, consulte a seção de classificações de configuração da versão do Amazon EMR que você usa. Por exemplo, Classificações de configuração do 6.2.0.

> **Note**
>
> O Guia de lançamento do Amazon EMR lista somente as ações de reconfiguração a partir das versões 5.32.0 e 6.2.0 do Amazon EMR.

Interrupção do serviço

O Amazon EMR segue um processo contínuo para reconfigurar as instâncias nos grupos de instâncias centrais e de tarefa. Somente 10% das instâncias em um grupo de instâncias são modificadas e reiniciadas por vez. Esse processo leva mais tempo para ser concluído, mas reduz a chance de possíveis falhas de aplicativo em um cluster em execução.

Para executar trabalhos do YARN durante a reinicialização do YARN, crie um cluster do Amazon EMR com múltiplos nós principais ou defina `yarn.resourcemanager.recovery.enabled` como `true` na classificação da configuração `yarn-site`. Para obter mais informações sobre o uso de múltiplos nós principais, consulte High availability YARN. ResourceManager

Validação da aplicação

O Amazon EMR verifica se cada aplicação do cluster está em execução após o processo de reinicialização da reconfiguração. Se alguma aplicação não estiver disponível, a operação geral de reconfiguração falhará. Se houver falha na reconfiguração de uma operação, o Amazon EMR reverterá os parâmetros de configuração para a versão de trabalho anterior.

> **Note**
>
> Para evitar falhas na reconfiguração, é recomendável instalar somente as aplicações que planeja usar no cluster. Também é recomendável verificar se todas as aplicações do cluster estão íntegras e em execução antes de enviar uma solicitação de reconfiguração.

Tipos de reconfiguração

É possível reconfigurar um grupo de instâncias de duas maneiras:

- Substituir. Método de reconfiguração padrão e o único disponível nas versões do Amazon EMR anteriores a 5.35.0 e 6.6.0. Esse método de reconfiguração substitui indiscriminadamente todo arquivo no cluster pelo conjunto de configurações recém-enviado. O método apaga as alterações nos arquivos de configuração feitas fora da API de reconfiguração.
- Mesclar. Método de reconfiguração compatível com o Amazon EMR 5.35.0 e 6.6.0 e versões posteriores, exceto no console do Amazon EMR, em que nenhuma versão é compatível. Esse método de reconfiguração mescla as configurações recém-enviadas com as configurações já existentes no cluster. Essa opção só adicionará ou modificará as novas configurações que você enviar. Ela preserva as configurações atuais.

> **Note**
>
> O Amazon EMR continua substituindo algumas configurações essenciais do Hadoop necessárias para garantir que o serviço seja executado corretamente.

Limitações

Ao reconfigurar um grupo de instâncias de um cluster em execução, considere estas limitações:

- Aplicações que não sejam do Yarn podem falhar durante a reinicialização ou causar problemas no cluster, sobretudo se as aplicações não estiverem configuradas corretamente. Ao se aproximarem demais do uso máximo de memória e CPU, os clusters poderão ter problemas após o processo de reinicialização. Isso se aplica principalmente ao grupo de instâncias principal.
- Não é possível enviar uma solicitação de reconfiguração quando o grupo de instâncias está sendo redimensionado. Se uma reconfiguração for iniciada enquanto um grupo de instâncias estiver sendo redimensionado, a reconfiguração não poderá ser iniciada até que o redimensionamento do grupo de instâncias tenha sido concluído e vice-versa.
- Depois de reconfigurar um grupo de instâncias, o Amazon EMR reiniciará as aplicações para permitir que as novas configurações entrem em vigor. Uma falha no trabalho ou outro comportamento inesperado do aplicativo poderão ocorrer se os aplicativos estiverem em uso durante a reconfiguração.
- Se houver falha na reconfiguração de um grupo de instâncias, o Amazon EMR reverterá os parâmetros de configuração para a versão anterior válida. Se também houver falha no processo

de reversão, você deverá enviar uma nova solicitação `ModifyInstanceGroup` para recuperar o grupo de instâncias a partir do estado `SUSPENDED`.

- As solicitações de reconfiguração para classificações de configuração do Phoenix são compatíveis com as versões 5.23.0 e posteriores do Amazon EMR, e não são compatíveis com a versão 5.21.0 ou 5.22.0 do Amazon EMR.
- Solicitações de reconfiguração para classificações de HBase configuração são compatíveis somente com o Amazon EMR 5.30.0 e versões posteriores, e não são compatíveis com as versões 5.23.0 a 5.29.0 do Amazon EMR.
- O Amazon EMR oferece suporte a solicitações de reconfiguração de aplicações em um cluster do Amazon EMR com múltiplos nós primários somente nas versões 5.27.0 e posteriores do Amazon EMR.
- Não há suporte para reconfigurar a classificação `hdfs-encryption-zones` ou qualquer uma das classificações de configuração em um cluster do Amazon EMR com múltiplos nós primários.
- O Amazon EMR atualmente não é compatível com determinadas solicitações de reconfiguração para o agendador de capacidade que exijam a reinicialização do YARN. ResourceManager Por exemplo, não é possível remover uma fila completamente.

Reconfigurar um grupo de instâncias no console

> **Note**
>
> O console do Amazon EMR não oferece suporte a reconfigurações do tipo Merge.

1. Abra o console do Amazon EMR em /emr https://console.aws.amazon.com
2. Na lista de clusters, m Nome, escolha o cluster ativo que você deseja reconfigurar.
3. Abra a página de detalhes do cluster e acesse a guia Configurações.
4. Na lista suspensa Filter (Filtro), selecione o grupo de instâncias que você deseja reconfigurar.
5. No menu suspenso Reconfigurar, escolha Editar na tabela ou Editar no arquivo JSON.

    - Editar na tabela: na tabela de classificações de configuração, edite a propriedade e o valor das configurações atuais ou escolha Adicionar configuração para fornecer a propriedade e o valor para outras classificações de configuração.

- Editar no arquivo JSON: insira a configuração diretamente no JSON ou use a sintaxe abreviada (demonstrada no texto sombreado). Caso contrário, forneça um URI do Amazon S3 para um arquivo com um objeto JSON `Configurations`.

> **Note**
>
> A coluna Origem na tabela de classificações de configuração indica se a configuração é fornecida quando você cria um cluster ou quando você especifica configurações adicionais para esse grupo de instâncias. Você pode editar as configurações para um grupo de instâncias de ambas as origens. Não é possível excluir as configurações iniciais do cluster, mas você pode substituí-las por um grupo de instâncias.
> Você também pode adicionar ou editar classificações de configuração aninhadas diretamente na tabela. Por exemplo, para fornecer uma sub-classificação de `export` adicional de `hadoop-env`, adicione uma classificação de configuração `hadoop.export` na tabela. Em seguida, forneça uma propriedade específica e um valor para essa classificação.

6. (Opcional) Selecione Apply this configuration to all active instance groups (Aplicar essa configuração a todos os grupos de instâncias ativas).
7. Salve as alterações.

Reconfigurar um grupo de instâncias usando a CLI

Você pode usar o comando modify-instance-groups para especificar configurações para cada grupo de instâncias em um cluster em execução.

> **Note**
>
> Nos exemplos a seguir, *`<j-2AL4XXXXXX5T9>`* substitua por seu ID de cluster e *`<ig-1xxxxxxx9>`* substitua pelo ID de seu grupo instâncias.

Example : substituir uma configuração para um grupo de instâncias

O exemplo a seguir referencia um arquivo de configuração JSON chamado `instanceGroups.json` para editar a propriedade do verificador de integridade do NodeManager disco YARN para um grupo de instâncias.

1. Prepare a classificação da configuração e salve-a como `instanceGroups.json` no mesmo diretório em que você executará o comando.

```
[
   {
      "InstanceGroupId":"<ig-1xxxxxxx9>",
      "Configurations":[
         {
            "Classification":"yarn-site",
            "Properties":{
               "yarn.nodemanager.disk-health-checker.enable":"true",
               "yarn.nodemanager.disk-health-checker.max-disk-utilization-per-disk-
percentage":"100.0"
            },
            "Configurations":[]
         }
      ]
   }
]
```

2. Execute o seguinte comando:

```
aws emr modify-instance-groups --cluster-id <j-2AL4XXXXXX5T9> \
--instance-groups file://instanceGroups.json
```

Example — Adicionar uma configuração a um grupo de instâncias

Se deseja adicionar uma configuração a um grupo de instâncias, você deverá incluir todas as configurações especificadas anteriormente para aquele grupo de instâncias em sua nova solicitação `ModifyInstanceGroup`. Caso contrário, as configurações especificadas anteriormente serão removidas.

O exemplo a seguir adiciona uma propriedade para o verificador de memória NodeManager virtual YARN. A configuração também inclui valores previamente especificados para o verificador de integridade do NodeManager disco YARN para que os valores não sejam substituídos.

1. Prepare o conteúdo a seguir em `instanceGroups.json` e salve-o no mesmo diretório onde você executará o comando.

```
[
   {
      "InstanceGroupId":"<ig-1xxxxxxx9>",
      "Configurations":[
         {
            "Classification":"yarn-site",
            "Properties":{
               "yarn.nodemanager.disk-health-checker.enable":"true",
               "yarn.nodemanager.disk-health-checker.max-disk-utilization-per-disk-
percentage":"100.0",
               "yarn.nodemanager.vmem-check-enabled":"true",
               "yarn.nodemanager.vmem-pmem-ratio":"3.0"
            },
            "Configurations":[]
         }
      ]
   }
]
```

2. Execute o seguinte comando:

```
aws emr modify-instance-groups --cluster-id <j-2AL4XXXXXX5T9> \
--instance-groups file://instanceGroups.json
```

Example — Adicionar uma configuração a um grupo de instâncias com uma reconfiguração do tipo Merge

Quando quiser usar o método padrão de reconfiguração Overwrite para adicionar uma configuração, é necessário incluir todas as configurações especificadas anteriormente para esse grupo de instâncias em sua nova solicitação `ModifyInstanceGroup`. Caso contrário, Overwrite removerá as configurações que você especificou anteriormente. Não é necessário fazer isso com a reconfiguração Merge. Em vez disso, garanta que sua solicitação inclua apenas as novas configurações.

O exemplo a seguir adiciona uma propriedade para o verificador de memória NodeManager virtual YARN. Como essa é uma reconfiguração do tipo Merge, ela não substitui os valores especificados anteriormente para o verificador de integridade do disco YARN. NodeManager

1. Prepare o conteúdo a seguir em `instanceGroups.json` e salve-o no mesmo diretório onde você executará o comando.

```
[
   {"InstanceGroupId":"<ig-1xxxxxxx9>",
    "ReconfigurationType" :"MERGE",
      "Configurations":[
         {"Classification":"yarn-site",
            "Properties":{
               "yarn.nodemanager.vmem-check-enabled":"true",
               "yarn.nodemanager.vmem-pmem-ratio":"3.0"
            },
            "Configurations":[]
         }
      ]
   }
]
```

2. Execute o seguinte comando:

```
aws emr modify-instance-groups --cluster-id <j-2AL4XXXXXX5T9> \
--instance-groups file://instanceGroups.json
```

Example : excluir uma configuração de um grupo de instâncias

Para excluir uma configuração para um grupo de instâncias, envie uma nova solicitação de reconfiguração que exclua a configuração anterior.

> **Note**
>
> Só é possível substituir a configuração inicial do cluster. Você não pode excluí-la.

Por exemplo, para excluir a configuração do verificador de integridade do NodeManager disco YARN do exemplo anterior, envie uma nova `instanceGroups.json` com o conteúdo a seguir.

```
[
   {
```

```
      "InstanceGroupId":"<ig-1xxxxxxx9>",
      "Configurations":[
         {
            "Classification":"yarn-site",
            "Properties":{
               "yarn.nodemanager.vmem-check-enabled":"true",
               "yarn.nodemanager.vmem-pmem-ratio":"3.0"
            },
            "Configurations":[]
         }
      ]
   }
]
```

Note

Para excluir todas as configurações na última solicitação de reconfiguração, envie uma solicitação de reconfiguração com uma matriz de configurações. Por exemplo,

```
[
   {
      "InstanceGroupId":"<ig-1xxxxxxx9>",
      "Configurations":[]
   }
]
```

Example — Reconfigurar e redimensionar um grupo de instâncias em uma solicitação

O JSON de exemplo a seguir demonstra como reconfigurar e redimensionar um grupo de instâncias em na mesma solicitação.

```
[
   {
      "InstanceGroupId":"<ig-1xxxxxxx9>",
      "InstanceCount":5,
      "EC2InstanceIdsToTerminate":["i-123"],
      "ForceShutdown":true,
      "ShrinkPolicy":{
```

```
         "DecommissionTimeout":10,
         "InstanceResizePolicy":{
            "InstancesToTerminate":["i-123"],
            "InstancesToProtect":["i-345"],
            "InstanceTerminationTimeout":20
         }
      },
      "Configurations":[
         {
            "Classification":"yarn-site",
            "Configurations":[],
            "Properties":{
               "yarn.nodemanager.disk-health-checker.enable":"true",
               "yarn.nodemanager.disk-health-checker.max-disk-utilization-per-disk-
percentage":"100.0"
            }
         }
      ]
   }
]
```

Reconfigurar um grupo de instâncias usando o SDK do Java

> **Note**
>
> Nos exemplos a seguir, *<j-2AL4XXXXXX5T9>* substitua por seu ID de cluster e *<ig-1xxxxxxx9>* substitua pelo ID de seu grupo instâncias.

O trecho de código a seguir fornece uma nova configuração para um grupo de instâncias usando o AWS SDK para Java.

```
AWSCredentials credentials = new BasicAWSCredentials("access-key", "secret-key");
AmazonElasticMapReduce emr = new AmazonElasticMapReduceClient(credentials);

Map<String,String> hiveProperties = new HashMap<String,String>();
hiveProperties.put("hive.join.emit.interval","1000");
hiveProperties.put("hive.merge.mapfiles","true");

Configuration configuration = new Configuration()
```

```
    .withClassification("hive-site")
    .withProperties(hiveProperties);

InstanceGroupModifyConfig igConfig = new InstanceGroupModifyConfig()
    .withInstanceGroupId("<ig-1xxxxxxx9>")
    .withReconfigurationType("MERGE");
    .withConfigurations(configuration);

ModifyInstanceGroupsRequest migRequest = new ModifyInstanceGroupsRequest()
    .withClusterId("<j-2AL4XXXXXX5T9>")
    .withInstanceGroups(igConfig);

emr.modifyInstanceGroups(migRequest);
```

O trecho de código a seguir exclui uma configuração especificada anteriormente para um grupo de instâncias fornecendo uma matriz vazia de configurações.

```
List<Configuration> configurations = new ArrayList<Configuration>();

InstanceGroupModifyConfig igConfig = new InstanceGroupModifyConfig()
    .withInstanceGroupId("<ig-1xxxxxxx9>")
    .withConfigurations(configurations);

ModifyInstanceGroupsRequest migRequest = new ModifyInstanceGroupsRequest()
    .withClusterId("<j-2AL4XXXXXX5T9>")
    .withInstanceGroups(igConfig);

emr.modifyInstanceGroups(migRequest);
```

Solucionar problemas de reconfiguração do grupo de instâncias

Se o processo de reconfiguração de um grupo de instâncias falhar, o Amazon EMR reverterá a reconfiguração e registrará em log uma mensagem de falha usando um evento da Amazon. CloudWatch O evento fornece um breve resumo da falha de reconfiguração. Ele lista as instâncias em que a reconfiguração falhou e as mensagens de falha correspondentes. Veja a seguir um exemplo de mensagem de falha.

```
The reconfiguration operation for instance group ig-1xxxxxxx9 in Amazon EMR
 cluster j-2AL4XXXXXX5T9 (ExampleClusterName)
```

```
failed at 2021-01-01 00:00 UTC and took 2 minutes to fail. Failed configuration version
 is example12345.
Failure message: Instance i-xxxxxxx1, i-xxxxxxx2, i-xxxxxxx3 failed with message "This
 is an example failure message".
```

Para coletar mais dados sobre uma falha de reconfiguração, verifique os logs de provisionamento do nó. Isso serve principalmente quando você recebe a mensagem a seguir.

```
i-xxxxxxx1 failed with message “Unable to complete transaction and some changes were
 applied.”
```

On the node

Acessar os logs de provisionamento do nó conectando-se a um nó

1. Use o SSH para se conectar ao nó em que a reconfiguração falhou. Para obter instruções, consulte Connect to your Linux instance no Amazon EC2 User Guide for Linux Instances.
2. Navegue até o diretório a seguir, que contém os arquivos de log de provisionamento do nó.

   ```
   /mnt/var/log/provision-node/
   ```

3. Abra o subdiretório `reports` e pesquise o relatório de provisionamento do nó para a reconfiguração. O `reports` diretório organiza os logs por número de versão de reconfiguração, identificador universal exclusivo (UUID), endereço IP da instância EC2 Amazon e carimbo de data/hora. Cada relatório é um arquivo YAML compactado contendo informações detalhadas sobre o processo de reconfiguração.

   Veja a seguir um exemplo de nome e caminho do arquivo de relatório.

   ```
   /reports/2/ca598xxx-cxxx-4xxx-bxxx-6dbxxxxxxxxx/ip-10-73-xxx-
   xxx.ec2.internal/202104061715.yaml.gz
   ```

4. Você pode examinar um relatório usando um visualizador de arquivos como `zless`, como no exemplo a seguir.

   ```
   zless 202104061715.yaml.gz
   ```

Amazon S3

Acessar logs de provisionamento do nó usando o Amazon S3

1. Faça login em AWS Management Console e abra o console da Amazon S3 em. https://console.aws.amazon.com/s3/
2. Abra o bucket do Amazon S3 especificado quando você configurou o cluster para arquivar arquivos de log.
3. Navegue até esta pasta, que contém os arquivos de log de provisionamento do nó:

   ```
   amzn-s3-demo-bucket/elasticmapreduce/<cluster id>/node/<instance id>/provision-node/
   ```

4. Abra a pasta `reports` e pesquise o relatório de provisionamento do nó para a reconfiguração. A `reports` pasta organiza os logs por número de versão de reconfiguração, identificador universal exclusivo (UUID), endereço IP da instância EC2 Amazon e carimbo de data/hora. Cada relatório é um arquivo YAML compactado contendo informações detalhadas sobre o processo de reconfiguração.

   Veja a seguir um exemplo de nome e caminho do arquivo de relatório.

   ```
   /reports/2/ca598xxx-cxxx-4xxx-bxxx-6dbxxxxxxxxx/ip-10-73-xxx-xxx.ec2.internal/202104061715.yaml.gz
   ```

5. Para visualizar um arquivo de log, você pode baixá-lo no Amazon S3 para a máquina local como arquivo de texto. Para obter instruções, consulte Fazer download de um objeto.

Cada arquivo de log contém um relatório detalhado de provisionamento para a reconfiguração associada. Para encontrar informações sobre mensagens de erro, pesquise o nível do log `err` de um relatório. O formato do relatório depende da versão do Amazon EMR no cluster.

O exemplo a seguir mostra informações de erro para versões do Amazon EMR anteriores a 5.32.0 e 6.2.0.

```
- !ruby/object:Puppet::Util::Log
      level: !ruby/sym err
      tags:
        - err
      message: "Example detailed error message."
      source: Puppet
```

```
      time: 2021-01-01 00:00:00.000000 +00:00
```

Em vez disso, as versões 5.32.0 e 6.2.0 e posteriores do Amazon EMR usam o formato a seguir.

```
- level: err
  message: 'Example detailed error message.'
  source: Puppet
  tags:
  - err
  time: '2021-01-01 00:00:00.000000 +00:00'
  file:
  line:
```

Armazenar dados de configuração sigilosos no AWS Secrets Manager

O Amazon EMR descreve e lista as operações de API que emitem dados de configuração personalizados (como `DescribeCluster` e `ListInstanceGroups`) em texto sem formatação. O Amazon EMR se integra ao AWS Secrets Manager para que você possa armazenar seus dados no Secrets Manager e usar o ARN do segredo em suas configurações. Dessa forma, você não passa dados de configuração sigilosos para o Amazon EMR em texto sem formatação nem os expõe ao ambiente externo. APIs Se você indicar que um par de chave-valor contém o ARN de um segredo armazenado no Secrets Manager, o Amazon EMR recuperará esse segredo ao enviar dados de configuração ao cluster. O Amazon EMR não envia a anotação ao usar a configuração externa APIs .

Criar um segredo

Para criar um segredo, siga as instruções em Criação de um segredo do AWS Secrets Manager no Guia do usuário do AWS Secrets Manager . Na Etapa 3, você deve escolher o campo Texto sem formatação para inserir o valor sigiloso.

Observe que, embora o Secrets Manager permita que o segredo contenha até 65536 bytes, o Amazon EMR limita a 1024 caracteres o tamanho combinado da chave de propriedade (excluindo a anotação) e o valor do segredo recuperado.

Conceder acesso ao Amazon EMR para recuperar o segredo

O Amazon EMR usa um perfil de serviço do IAM para provisionar e gerenciar clusters para você. O perfil de serviço do Amazon EMR define as ações permitidas para o Amazon EMR durante o

provisionamento de recursos e a execução de tarefas no nível de serviço que não são executadas no contexto de uma instância da Amazon em execução em um cluster. EC2 Para obter mais informações sobre perfis de serviço, consulte Service role for Amazon EMR (EMR role) e Customize IAM roles.

Para permitir que o Amazon EMR recupere o valor do segredo do Secrets Manager, adicione a seguinte instrução de política ao perfil do Amazon EMR ao iniciar o cluster.

```
{
   "Sid":"AllowSecretsRetrieval",
   "Effect":"Allow",
   "Action":"secretsmanager:GetSecretValue",
   "Resource":[
      "arn:aws:secretsmanager:<region>:<aws-account-id>:secret:<secret-name>"
   ]
}
```

Ao criar o segredo com uma chave do gerenciada pelo cliente AWS KMS key, você também deverá adicionar `kms:Decrypt` permissão ao perfil do Amazon EMR para a chave utilizada. Para obter mais informações, consulte Controle de acesso e autenticação para o AWS Secrets Manager no Manual do usuário do AWS Secrets Manager .

Usar o segredo em uma classificação de configuração

Você pode adicionar a anotação `EMR.secret@` a qualquer propriedade de configuração para indicar que o par de chave-valor contém um ARN de um segredo armazenado no Secrets Manager.

O exemplo a seguir mostra como fornecer um ARN de segredo em uma classificação de configuração:

```
{
   "Classification":"core-site",
   "Properties":{
      "presto.s3.access-key":"<sensitive-access-key>",
      "EMR.secret@presto.s3.secret-key":"arn:aws:secretsmanager:<region>:<aws-account-
id>:secret:<secret-name>"
   }
}
```

Quando você cria o cluster e envia a configuração anotada, o Amazon EMR valida as propriedades da configuração. Se a configuração for válida, o Amazon EMR removerá a anotação da configuração

e recuperará o segredo do Secrets Manager para criar a configuração real antes de aplicá-la ao cluster:

```
{
   "Classification":"core-site",
   "Properties":{
      "presto.s3.access-key":"<sensitive-access-key>",
      "presto.s3.secret-key":"<my-secret-key-retrieved-from-Secrets-Manager>"
   }
}
```

Quando você chama uma ação como `DescribeCluster`, o Amazon EMR retorna a configuração atual da aplicação no cluster. Se uma propriedade de configuração da aplicação estiver marcada como contendo um ARN de segredo, a configuração da aplicação retornada pela chamada `DescribeCluster` conterá o ARN e não o valor do segredo. Isso garante que o valor do segredo seja visível somente no cluster:

```
{
   "Classification":"core-site",
   "Properties":{
      "presto.s3.access-key":"<sensitive-access-key>",
      "presto.s3.secret-key":"arn:aws:secretsmanager:<region>:<aws-account-
id>:secret:<secret-name>"
   }
}
```

Atualizar o valor do segredo

O Amazon EMR recupera o valor do segredo de uma configuração anotada sempre que o grupo de instâncias anexado estiver iniciando, reconfigurando ou redimensionando. É possível usar o Secrets Manager para modificar o valor de um segredo utilizado na configuração de um cluster em execução. Ao fazer isso, você pode enviar uma solicitação de reconfiguração para cada grupo de instâncias que deve receber o valor atualizado. Para obter mais informações sobre como reconfigurar um grupo de instâncias e o que considerar ao fazer isso, consulte Reconfigurar um grupo de instâncias de um cluster em execução.

Configurar aplicações para usar uma Máquina Virtual Java específica

As versões do Amazon EMR têm diferentes versões Máquina Virtual Java (JVM) padrão. Esta página explica a compatibilidade da JVM para diferentes versões e aplicações.

Considerações

Para obter informações sobre as versões do Java compatíveis com aplicações, consulte as páginas das aplicações no Guia de lançamento do Amazon EMR.

- O Amazon EMR só oferece suporte à execução de uma versão de runtime em um cluster e não oferece suporte à execução de nós ou aplicações diferentes em outras versões de runtime no mesmo cluster.
- Para o Amazon EMR 7.x, a Máquina Virtual Java (JVM) padrão é Java 17 para aplicações compatíveis com Java 17, com exceção do Apache Livy. Para obter mais informações sobre as versões do JDK compatíveis com aplicações, consulte a página de lançamento correspondente no Guia de lançamento do Amazon EMR.
- A partir do Amazon EMR 7.1.0, o Flink oferece suporte e está configurado para Java 17 por padrão. Para usar uma versão diferente do runtime do Java, substitua as configurações em `flink-conf`. Para obter mais informações sobre como configurar o Flink para usar o Java 8 ou o Java 11, consulte Configure Flink to run with Java 11.
- Para as séries 5.x e 6.x do Amazon EMR, a Máquina Virtual Java (JVM) padrão é Java 8.
  - Para as versões 6.12.0 e posteriores do Amazon EMR, algumas aplicações também oferecem suporte ao Java 11 e 17.
  - Para as versões 6.9.0 e posteriores do Amazon EMR, o Trino oferece suporte ao Java 17 como padrão. Para obter mais informações sobre o Java 17 com Trino, consulte Trino updates to Java 17 no blog do Trino.

Lembre-se das seguintes considerações específicas da aplicação ao escolher sua versão de runtime:

Notas de configuração Java específicas da aplicação

Aplicação	Notas sobre configurações do Java
Spark	Para executar o Spark com uma versão Java que não seja padrão, é necessário configura r o Spark e o Hadoop. Para obter exemplos, consulte Substituir a JVM. • Configure `JAVA_HOME` em `spark-env` para atualizar o runtime Java dos processos da instância primária. Por exemplo, spark-submit, spark-shell e Spark History Server. • Modifique a configuração do Hadoop para atualizar o runtime Java dos executores do Spark e do YARN `ApplicationMaster`
Spark RAPIDS	É possível executar o RAPIDS com a versão Java configurada para o Spark.
Iceberg	Você pode executar o Iceberg com a versão Java configurada da aplicação que o está usando.
Delta	Você pode executar o Delta com a versão Java configurada da aplicação que o está usando.
Hudi	Você pode executar o Hudi com a versão Java configurada da aplicação que o está usando.
Hadoop	Para atualizar a JVM para o Hadoop, modifique `hadoop-env` . Para obter exemplos, consulte Substituir a JVM.
Hive	Para definir a versão Java como 11 ou 17 para Hive, defina a configuração da JVM do Hadoop para a versão do Java que você deseja usar.

Aplicação	Notas sobre configurações do Java
HBase	Para atualizar a JVM para HBase, modifique . `hbase-env` Por padrão, o Amazon EMR define a HBase JVM segundo a configuração da JVM para o Hadoop, a menos que você substitua as configurações em. `hbase-env` Para obter exemplos, consulte Substituir a JVM.
Flink	Para atualizar a JVM para Flink, modifique `flink-conf` . Por padrão, o Amazon EMR define a JVM do Flink segundo a configuração da JVM para o Hadoop, a menos que você substitua as configurações em `flink-con f` . Para obter mais informações, consulte Configurar o Flink para ser executado com o Java 11.
Oozie	Para configurar o Oozie para ser executado no Java 11 ou 17, configure o Oozie Server, o Oozie LauncherAM Launcher AM e altere as configurações de trabalho e executáve is do lado do cliente. Também é possível configurar `EmbeddedOozieServer` para ser executado no Java 17. Para obter mais informações, consulte Configurar a versão do Java para Oozie.
Pig	O Pig é compatível apenas com Java 8. Não é possível usar o Java 11 ou 17 com o Hadoop e executar o Pig no mesmo cluster.

Substituir a JVM

Para substituir a configuração da JVM para uma versão do Amazon EMR (por exemplo, para usar o Java 17 com um cluster que usa o Amazon EMR versão 6.12.0) forneça a configuração `JAVA_HOME`

para sua classificação de ambiente, que é *application*`-env` para todas as aplicações, exceto o Flink. Para o Flink, a classificação do ambiente é `flink-conf`. Para obter as etapas para configurar o runtime Java com o Flink, consulte Configurar o Flink para ser executado com o Java 11.

Tópicos

- Substituir a configuração da JVM usando o Apache Spark
- Substituir a configuração da JVM usando o Apache HBase
- Substituir a configuração da JVM usando o Apache Hadoop e o Hive

Substituir a configuração da JVM usando o Apache Spark

Ao usar o Spark com o Amazon EMR versões 6.12 e posteriores, você pode configurar o ambiente para que os executores usem Java 11 ou 17. E quando você usa o Spark com versões do Amazon EMR anteriores a 5.x e grava um driver para envio no modo de cluster, o driver usa o Java 7. No entanto, você pode configurar o ambiente para garantir que os executores usem o Java 8.

Para substituir a JVM do Spark, você deve definir as classificações do Hadoop e do Spark.

```
{
"Classification": "hadoop-env",
        "Configurations": [
            {
"Classification": "export",
                "Configurations": [],
                "Properties": {
"JAVA_HOME": "/usr/lib/jvm/java-1.8.0"
                }
            }
        ],
        "Properties": {}
    },
    {
"Classification": "spark-env",
        "Configurations": [
            {
"Classification": "export",
                "Configurations": [],
                "Properties": {
"JAVA_HOME": "/usr/lib/jvm/java-1.8.0"
                }
```

```
            }
        ],
        "Properties": {}
    }
```

O exemplo a seguir mostra como adicionar os parâmetros de configuração necessários para o EMR 7.0.0+ para garantir o uso consistente da versão Java em todos os componentes.

```
[
  {
    "Classification": "spark-defaults",
    "Properties": {
      "spark.executorEnv.JAVA_HOME": "/usr/lib/jvm/java-1.8.0",
      "spark.yarn.appMasterEnv.JAVA_HOME": "/usr/lib/jvm/java-1.8.0"
    }
  },
  {
    "Classification": "hadoop-env",
    "Configurations": [
      {
        "Classification": "export",
        "Configurations": [],
        "Properties": {
          "JAVA_HOME": "/usr/lib/jvm/java-1.8.0"
        }
      }
    ],
    "Properties": {}
  },
  {
    "Classification": "spark-env",
    "Configurations": [
      {
        "Classification": "export",
        "Configurations": [],
        "Properties": {
          "JAVA_HOME": "/usr/lib/jvm/java-1.8.0"
        }
      }
    ],
    "Properties": {}
  }
]
```

Substituir a configuração da JVM usando o Apache HBase

Para configurar HBase para usar o Java 11, você pode definir a configuração a seguir ao iniciar o cluster.

```
[
  {
    "Classification": "hbase-env",
    "Configurations": [
      {
        "Classification": "export",
        "Configurations": [],
        "Properties": {
          "JAVA_HOME": "/usr/lib/jvm/jre-11"
        }
      }
    ],
    "Properties": {}
  }
]
```

Substituir a configuração da JVM usando o Apache Hadoop e o Hive

O exemplo a seguir mostra como definir a JVM para a versão 17 para Hadoop e Hive.

```
[
    {
        "Classification": "hadoop-env",
            "Configurations": [
                {
                    "Classification": "export",
                    "Configurations": [],
                    "Properties": {
                        "JAVA_HOME": "/usr/lib/jvm/jre-17"
                    }
                }
        ],
        "Properties": {}
    }
]
```

Portas de serviço

A seguir estão as portas de serviço YARN e HDFS. Essas configurações refletem os padrões do Hadoop. Outros serviços de aplicativos são hospedados em portas padrão, a menos que documentado de outra forma. Para obter mais informações, consulte a documentação de projeto do aplicativo.

Configurações de portas para o YARN e o HDFS

Configuração	Nome do host/Porta
`fs.default.name`	(padrão) (`hdfs://`*`emrDeterminedIP`* `:8020`)
`dfs.datanode.address`	(padrão) (`0.0.0.0:50010`)
`dfs.datanode.http.address`	(padrão) (`0.0.0.0:50075`)
`dfs.datanode.https.address`	(padrão) (`0.0.0.0:50475`)
`dfs.datanode.ipc.address`	(padrão) (`0.0.0.0:50020`)
`dfs.http.address`	(padrão) (`0.0.0.0:50070`)
`dfs.https.address`	(padrão) (`0.0.0.0:50470`)
`dfs.secondary.http.address`	(padrão) (`0.0.0.0:50090`)
`yarn.nodemanager.address`	(padrão) (`${yarn.nodemanager.hostname}:0`)
`yarn.nodemanager.localizer.address`	(padrão) (`${yarn.nodemanager.hostname}:8040`)
`yarn.nodemanager.webapp.address`	(padrão) (`${yarn.nodemanager.hostname}:8042`)
`yarn.resourcemanager.address`	(padrão) (`${yarn.resourcemanager.hostname}:8032`)

Configuração	Nome do host/Porta
yarn.resourcemanager.admin.address	(padrão) (${yarn.resourceman ager.hostname}:8033)
yarn.resourcemanager.resource-tracker.address	(padrão) (${yarn.resourceman ager.hostname}:8031)
yarn.resourcemanager.schedu ler.address	(padrão) (${yarn.resourceman ager.hostname}:8030)
yarn.resourcemanager.webapp .address	(padrão) (${yarn.resourceman ager.hostname}:8088)
yarn.web-proxy.address	(padrão) (no-value)
yarn.resourcemanager.hostname	*emrDeterminedIP*

> **Note**
>
> O termo *emrDeterminedIP* é um endereço IP gerado pelo ambiente de gerenciamento do Amazon EMR. Na versão mais recente, essa convenção foi eliminada, com exceção para as configurações yarn.resourcemanager.hostname e fs.default.name.

Usuários do aplicativo

Os aplicativos executarão processos como seus próprios usuários. Por exemplo, o Hive é JVMs executado como usuáriohive, MapReduce JVMs executado como mapred e assim por diante. Isso é demonstrado no exemplo a seguir de status de processo.

```
USER       PID %CPU %MEM    VSZ   RSS TTY      STAT START   TIME COMMAND
hive      6452  0.2  0.7 853684 218520 ?       Sl   16:32   0:13 /usr/lib/jvm/
java-openjdk/bin/java -Xmx256m -Dhive.log.dir=/var/log/hive -Dhive.log.file=hive-
metastore.log -Dhive.log.threshold=INFO -Dhadoop.log.dir=/usr/lib/hadoop
hive      6557  0.2  0.6 849508 202396 ?       Sl   16:32   0:09 /usr/lib/jvm/java-
openjdk/bin/java -Xmx256m -Dhive.log.dir=/var/log/hive -Dhive.log.file=hive-server2.log
 -Dhive.log.threshold=INFO -Dhadoop.log.dir=/usr/lib/hadoop/l
```

```
hbase     6716  0.1  1.0 1755516 336600 ?      Sl   Jun21   2:20 /usr/lib/jvm/java-
openjdk/bin/java -Dproc_master -XX:OnOutOfMemoryError=kill -9 %p -Xmx1024m -ea -XX:
+UseConcMarkSweepGC -XX:+CMSIncrementalMode -Dhbase.log.dir=/var/
hbase     6871  0.0  0.7 1672196 237648 ?      Sl   Jun21   0:46 /usr/lib/jvm/java-
openjdk/bin/java -Dproc_thrift -XX:OnOutOfMemoryError=kill -9 %p -Xmx1024m -ea -XX:
+UseConcMarkSweepGC -XX:+CMSIncrementalMode -Dhbase.log.dir=/var/
hdfs      7491  0.4  1.0 1719476 309820 ?      Sl   16:32   0:22 /usr/lib/jvm/java-
openjdk/bin/java -Dproc_namenode -Xmx1000m -Dhadoop.log.dir=/var/log/hadoop-hdfs -
Dhadoop.log.file=hadoop-hdfs-namenode-ip-10-71-203-213.log -Dhadoo
yarn      8524  0.1  0.6 1626164 211300 ?      Sl   16:33   0:05 /usr/lib/jvm/java-
openjdk/bin/java -Dproc_proxyserver -Xmx1000m -Dhadoop.log.dir=/var/log/hadoop-yarn -
Dyarn.log.dir=/var/log/hadoop-yarn -Dhadoop.log.file=yarn-yarn-
yarn      8646  1.0  1.2 1876916 385308 ?      Sl   16:33   0:46 /usr/lib/jvm/java-
openjdk/bin/java -Dproc_resourcemanager -Xmx1000m -Dhadoop.log.dir=/var/log/hadoop-yarn
 -Dyarn.log.dir=/var/log/hadoop-yarn -Dhadoop.log.file=yarn-y
mapred    9265  0.2  0.8 1666628 260484 ?      Sl   16:33   0:12 /usr/lib/jvm/java-
openjdk/bin/java -Dproc_historyserver -Xmx1000m -Dhadoop.log.dir=/usr/lib/hadoop/logs -
Dhadoop.log.file=hadoop.log -Dhadoop.home.dir=/usr/lib/hadoop
```

Verificar dependências usando o repositório de artefatos do Amazon EMR

Você pode usar o repositório de artefatos do Amazon EMR para compilar o código de trabalho do Apache Hive e do Apache Hadoop relacionado às versões exatas de bibliotecas e dependências disponíveis com versões de específicas do Amazon EMR, a partir da versão 5.18.0 do Amazon EMR. A compilação de artefatos do Amazon EMR no repositório ajuda a evitar problemas do caminho de classe do runtime, garantindo que as versões das bibliotecas nas quais o trabalho é criado sejam exatamente as mesmas versões fornecidas no runtime no cluster. Atualmente, os artefatos do Amazon EMR estão disponíveis somente para compilações do Maven.

Para acessar o repositório de artefatos, adicione o URL do repositório ao arquivo de configurações Maven ou ao arquivo de configurações `pom.xml` de um projeto específico. Em seguida, você pode especificar as dependências na configuração de seu projeto. Para as versões de dependência, use a versão listada em Versões do componente para a versão desejada em Versões de lançamento 5.x do Amazon EMR. Por exemplo, as versões de componente para a versão do Amazon EMR mais recente está disponível em the section called “Versões de componente 5.36.2”. Se um artefato para o seu projeto não estiver listado em Versões do componente, especifique a versão que está listada para o Hive e o Hadoop nessa versão. Por exemplo, para componentes do Hadoop na versão 5.18.0 do Amazon EMR, a versão é `2.8.4-amzn-1`.

O URL do repositório de artefatos tem a seguinte sintaxe:

```
https://s3-endpoint/region-ID-emr-artifacts/emr-release-label/repos/maven/
```

- *`s3-endpoint`* é o endpoint do Amazon Simple Storage Service (Amazon S3) da região para o repositório *`region-ID`* e é a região correspondente. Por exemplo, `s3.us-west-1.amazonaws.com` e `us-west-1`. Para obter mais informações, consulte os endpoints do Amazon S3 na Referência geral da Amazon Web Services. Não há diferença nos artefatos entre as regiões, portanto, você pode especificar a região mais conveniente para o seu ambiente de desenvolvimento.
- *`emr-release-label`* é a etiqueta de lançamento do cluster Amazon EMR que executará seu código. Os rótulos de versão estão no formato `emr-x.x.x`, como `emr-5.36.2`. Uma série de versões do EMR pode incluir várias versões. Por exemplo, se você estiver usando a versão 5.24.1 do EMR, use o primeiro rótulo de versão do EMR na série 5.24, `emr-5.24.0`, no URL do repositório de artefatos:

```
https://s3-endpoint/region-ID-emr-artifacts/emr-5.24.0/repos/maven/
```

Example Configuração para pom.xml do Maven

O exemplo pom.xml a seguir configura um projeto Maven para compilar os artefatos do Apache Hadoop e do Apache Hive do emr-5.18.0 usando o repositório de artefatos em us-west-1. As versões de snapshot não estão disponíveis no repositório de artefatos. Portanto, os snapshots estão desabilitados no pom.xml. As elipses (`...`) no exemplo abaixo indicam a omissão de outros parâmetros de configuração. Não os copie em seu projeto Maven.

```
<project>
 ...
 <repositories>
  ...
  <repository>
   <id>emr-5.18.0-artifacts</id>
   <name>EMR 5.18.0 Releases Repository</name>
   <releases>
    <enabled>true</enabled>
   </releases>
   <snapshots>
    <enabled>false</enabled>
   </snapshots>
   <url>https://s3.us-west-1.amazonaws.com/us-west-1-emr-artifacts/emr-5.18.0/repos/
maven/</url>
  </repository>
  ...
 </repositories>
 ...
 <dependencies>
  ...
  <dependency>
   <groupId>org.apache.hive</groupId>
   <artifactId>hive-exec</artifactId>
   <version>2.3.3-amzn-2</version>
  </dependency>
  <dependency>
   <groupId>org.apache.hadoop</groupId>
   <artifactId>hadoop-common</artifactId>
   <version>2.8.4-amzn-1</version>
  </dependency>
```

```
  ...
 </dependencies>

</project>
```

Sistema de arquivos do EMR (EMRFS)

O EMR File System (EMRFS) é uma implementação do HDFS que todos os clusters do Amazon EMR usam para leitura e gravação de arquivos regulares do Amazon EMR diretamente no Amazon S3. O EMRFS oferece a conveniência de armazenar dados persistentes no Amazon S3 para serem usados com o Hadoop, fornecendo também atributos como criptografia de dados.

A criptografia de dados permite criptografar objetos que o EMRFS grava no Amazon S3 e permite que o EMRFS trabalhe com objetos criptografados no Amazon S3. Se estiver usando o Amazon EMR 4.8.0 ou versões posteriores, você poderá usar as configurações de segurança para configurar a criptografia para objetos do EMRFS no Amazon S3, junto com outras configurações de criptografia. Para obter mais informações, consulte Encryption options. Ao usar uma versão anterior do Amazon EMR, você pode definir as configurações de criptografia manualmente. Para obter mais informações, consulte Especificar a criptografia do Amazon S3 usando propriedades do EMRFS.

O Amazon S3 oferece uma consistência avançada de leitura após gravação para todas as operações GET, PUT e LIST em todas as Regiões da AWS. Com isso, o que você grava usando o EMRFS é o que você lê no Amazon S3, sem afetar a performance. Para obter mais informações, consulte Modelo de consistência de dados do Amazon S3.

Ao usar o Amazon EMR versão 5.10.0 ou posterior, você pode usar diferentes perfis do IAM para solicitações do EMRFS ao Amazon S3 com base em usuários do cluster, grupos ou no local dos dados do EMRFS no Amazon S3. Para obter mais informações, consulte Configure IAM roles for EMRFS requests to Amazon S3.

> **Warning**
>
> Antes de ativar a execução especulativa em clusters do Amazon EMR que executam trabalhos do Apache Spark, analise as informações a seguir.
> O EMRFS inclui o confirmador otimizado para EMRFS S3, uma OutputCommitter implementação que foi otimizada para gravar arquivos no Amazon S3 ao usar o EMRFS. Se ativar o atributo de execução especulativa do Apache Spark com aplicações que gravam dados no Amazon S3 e não usam o confirmador otimizado para EMRFS S3, você poderá encontrar problemas de correção de dados descritos em SPARK-10063. Isso poderá ocorrer se você estiver usando versões do Amazon EMR anteriores à versão 5.19 do Amazon EMR ou se estiver gravando arquivos no Amazon S3 com formatos como ORC e CSV. Esses formatos não são compatíveis com o confirmador otimizado para EMRFS S3. Para obter uma

lista completa dos requisitos para usar o confirmador otimizado para EMRFS S3, consulte Requirements for the EMRFS S3-optimized committer.
A gravação direta do EMRFS geralmente é usada quando não há suporte para o confirmador otimizado para EMRFS S3, como ao gravar:

- Um formato de saída diferente de Parquet, como ORC ou texto.
- Arquivos Hadoop usando a API Spark RDD.
- Parquet usando Hive SerDe. Consulte Hive metastore Parquet table conversion.

Não se usa a gravação direta do EMRFS nos seguintes cenários:

- Quando o committer otimizado para EMRFS S3 não está habilitado. Consulte Requirements for the EMRFS S3-optimized committer.
- Ao escrever partições dinâmicas com partitionOverwriteMode configurado como dinâmico.
- Ao gravar em locais de partição personalizados, como locais fora de conformidade com a convenção de localização de partições padrão do Hive.
- Ao usar sistemas de arquivos que não são do EMRFS, como gravar no HDFS ou usar o sistema de arquivos S3A.

Para determinar se sua aplicação usa gravação direta no Amazon EMR 5.14.0 ou versão posterior, habilite o registro em log do Spark INFO. Se houver uma linha de log contendo o texto “Gravação direta: HABILITADA” nos logs do driver do Spark ou nos logs do contêiner do executor do Spark, sua aplicação Spark gravou usando gravação direta.
Por padrão, a execução especulativa é habilitada `OFF` na Amazon EMRclusters. É altamente recomendável não ativar a execução especulativa se estas duas condições forem verdadeiras:

- Você está gravando dados no Amazon S3.
- Os dados são gravados em um formato que não seja Apache Parquet ou no formato Apache Parquet sem usar o confirmador otimizado para EMRFS S3.

Se você ativar a execução especulativa do Spark e gravar dados no Amazon S3 usando a gravação direta do EMRFS, poderá ocorrer perda de dados intermitente. Quando você grava dados no HDFS, ou grava dados no Parquet usando o confirmador otimizado para EMRFS S3, o Amazon EMR não usa gravação direta, e esse problema não ocorre.

Se você precisar gravar dados em formatos que usem a gravação direta do EMRFS do Spark para o Amazon S3 e usar a execução especulativa, recomenda-se gravar no HDFS e depois transferir os arquivos de saída para o Amazon S3 usando S3DistCP.

Tópicos

- Consistent view
- Autorizar acesso aos dados do EMRFS no Amazon S3
- Gerenciar o AWS Security Token Service endpoint do padrão
- Especificar a criptografia do Amazon S3 usando propriedades do EMRFS

Consistent view

Warning

Em 1.º de junho de 2023, a visualização consistente do EMRFS chegará ao fim do suporte padrão para futuras versões do Amazon EMR. A visualização consistente do EMRFS continuará funcionando para as versões atuais.

Com a versão de read-after-write consistência avançada do Amazon S3 em 1.º de dezembro de 2020, você não precisa mais usar a visualização consistente do EMRFS (EMRFS CV) com seus clusters do Amazon EMR. A EMRFS CV é um atributo opcional que permite que os clusters do Amazon EMR verifiquem a lista e a consistência dos objetos do read-after-write Amazon S3. Quando você cria um cluster e a EMRFS CV está ativada, o Amazon EMR cria um banco de dados do Amazon DynamoDB para armazenar metadados de objetos que ele usa para rastrear a consistência e a lista dos objetos do S3. read-after-write Agora é possível desativar a EMRFS CV e excluir o banco de dados do DynamoDB que ela usa para não acumular custos adicionais. Os procedimentos a seguir explicam como verificar o atributo CV, desativá-lo e excluir o banco de dados do DynamoDB usado pelo atributo.

Verificar se você está usando o recurso EMRFS CV

1. Navegue até a guia Configuração. Se o cluster tem a configuração a seguir, ele usa a EMRFS CV.

```
Classification=emrfs-site,Property=fs.s3.consistent,Value=true
```

2. Como alternativa, use o AWS CLI para descrever seu cluster com a `describe-cluster`API. Se a saída contém `fs.s3.consistent: true`, o cluster usa a EMRFS CV.

Desativar a EMRFS CV nos seus clusters do Amazon EMR

Para desativar o atributo EMRFS CV, use uma destas três opções. É necessário testar essas opções em seu ambiente de teste antes de aplicá-las aos ambientes de produção.

1. Para interromper o cluster existente e iniciar um novo cluster sem as opções da EMRFS CV.
   a. Antes de interromper o cluster, faça backup dos dados e notifique os usuários.
   b. Para interromper o cluster, siga as instruções apresentadas em Terminate a cluster.
   c. Se você usar o console do Amazon EMR para criar um novo cluster, navegue até Opções avançadas. Na seção Editar configurações do software, desmarque a opção para ativar a EMRFS CV. Se a caixa de seleção Visualização consistente do EMRFS estiver disponível, mantenha-a desmarcada.
   d. Se você usar AWS CLI a para criar um novo cluster com a `create-cluster`API, não use a `--emrfs` opção, que ativa a EMRFS CV.
   e. Se você usar um SDK ou AWS CloudFormation o para criar um novo cluster, não use nenhuma das configurações listadas em Configure consistent view.
2. Clonar um cluster e remover a EMRFS CV
   a. No console do Amazon EMR, escolha o cluster que está usando a EMRFS CV.
   b. Na parte superior da página Detalhes do cluster, escolha Clonar.
   c. Escolha Anterior e navegue até a Etapa 1: software e etapas.
   d. Em Editar configurações de software, remova a EMRFS CV. Em Editar configuração, exclua as seguintes configurações na classificação `emrfs-site`. Se você estiver carregando JSON de um bucket do S3, será necessário modificar o objeto do S3.

```
[
 {"classification":
  "emrfs-site",
  "properties": {
   "fs.s3.consistent.retryPeriodSeconds":"10",
```

```
    "fs.s3.consistent":"true",
    "fs.s3.consistent.retryCount":"5",
    "fs.s3.consistent.metadata.tableName":"EmrFSMetadata"
   }
  }
]
```

3. Remover a EMRFS CV de um cluster que usa grupos de instâncias

   a. Use o comando a seguir para verificar se apenas um cluster do EMR usa a tabela do DynamoDB associada à EMRFS CV ou se vários clusters compartilham a tabela. O nome da tabela é especificado em `fs.s3.consistent.metadata.tableName`, conforme descrito em Configure consistent view. O nome da tabela padrão usado pela EMRFS CV é `EmrFSMetadata`.

   ```
   aws emr describe-cluster --cluster-id j-XXXXX | grep
    fs.s3.consistent.metadata.tableName
   ```

   b. Se o cluster não compartilhar o banco de dados do DynamoDB com outro cluster, use o comando a seguir para reconfigurar o cluster e desativar a EMRFS CV. Para obter mais informações, consulte Reconfigure an instance group in a running cluster.

   ```
   aws emr modify-instance-groups --cli-input-json file://disable-emrfs-1.json
   ```

   Esse comando abrirá o arquivo que você deseja modificar. Modifique o arquivo com as configurações a seguir.

   ```
   {
    "ClusterId": "j-xxxx",
    "InstanceGroups": [
     {
      "InstanceGroupId": "ig-xxxx",
      "Configurations": [
       {
        "Classification": "emrfs-site",
        "Properties": {
         "fs.s3.consistent": "false"
        },
        "Configurations": []
       }
      ]
   ```

```
  }
 ]
}
```

c. Se o cluster compartilhar a tabela do DynamoDB com outro cluster, desative a EMRFS CV em todos os clusters ao mesmo tempo em que nenhum cluster modifique nenhum objeto no local compartilhado do S3.

Excluir recursos do Amazon DynamoDB associados à EMRFS CV

Após remover a EMRFS CV dos clusters do Amazon EMR, exclua os recursos do DynamoDB associados à EMRFS CV. Enquanto não fizer isso, você continuará incorrendo nas cobranças do DynamoDB associadas à EMRFS CV.

1. Verifique as CloudWatch métricas da sua tabela do DynamoDB e confirme se a tabela não está sendo usada por nenhum cluster.
2. Exclua a tabela do DynamoDB.

```
aws dynamodb delete-table --table-name <your-table-name>
```

Excluir recursos do Amazon SQS associados à EMRFS CV

1. Se você configurou o cluster para enviar notificações de inconsistência para o Amazon SQS, será possível excluir todas as filas do SQS.
2. Encontre o nome da fila do Amazon SQS especificado em `fs.s3.consistent.notification.SQS.queueName`, conforme descrito em Configure consistent view. O formato padrão do nome da fila é `EMRFS-Inconsistency-<j-cluster ID>`.

```
aws sqs list-queues | grep 'EMRFS-Inconsistency'
aws sqs delete-queue -queue-url <your-queue-url>
```

Parar de usar a CLI do EMRFS

- A CLI do EMRFS gerencia os metadados que a EMRFS CV gera. À medida que o suporte padrão para a EMRFS CV chegar ao fim nas futuras versões do Amazon EMR, o suporte para a CLI do EMRFS também terminará.

Tópicos

- Habilitar a visualização consistente
- Noções básicas sobre como a visualização consistente do EMRFS rastreia objetos no Amazon S3
- Lógica de novas tentativas
- Metadados da visualização consistente do EMRFS
- Configurar notificações de consistência para CloudWatch o Amazon SQS
- Configurar a visualização consistente
- Referência de comando da CLI do EMRFS

Habilitar a visualização consistente

É possível habilitar a criptografia do lado do servidor do Amazon S3 ou a visualização consistente do EMRFS usando o, a ou a AWS Management Console classificação de configuração. AWS CLI`emrfs-site`

Para configurar a exibição consistente usando o console

1. Navegue até o novo console do Amazon EMR e selecione Alternar para o console antigo na navegação lateral. Para obter mais informações sobre o que esperar ao alternar para o console antigo, consulte Usar o console antigo.
2. Escolha Create cluster (Criar cluster), Go to advanced options (Ir para opções avançadas).
3. Escolha as configurações de Step 1: Software and Steps (Etapa 1: Software e etapas) e Step 2: Hardware (Etapa 2: Hardware).
4. Em Step 3: General Cluster Settings (Etapa 3: Configurações gerais de cluster), em Additional Options (Opções adicionais), escolha EMRFS consistent view (Visualização consistente de EMRFS).
5. Para EMRFS Metadata store (Armazenamento de metadados EMRFS), digite o nome do seu repositório de metadados. O valor padrão é **`EmrFSMetadata`**. Se a FSMetadata tabela Emr não existir, ela será criada para você no DynamoDB.

> **Note**
>
> O Amazon EMR não remove automaticamente os metadados do EMRFS do DynamoDB quando o cluster é terminado.

6. Para Number of retries (Número de novas tentativas), digite um valor inteiro. Caso uma inconsistência seja detectada, o EMRFS tenta chamar o Amazon S3 esse número de vezes. O valor padrão é **5**.
7. Para Retry period (in seconds) [Período de repetição (em segundos)], digite um valor inteiro. Esse é o período que o EMRFS aguarda entre as tentativas. O valor padrão é **10**.

> **Note**
>
> Tentativas subsequentes usam um recuo exponencial.

Para iniciar um cluster com visualização consistente habilitadas usando a AWS CLI

Recomendamos instalar a versão atual do AWS CLI. Para baixar a versão mais recente, consulte https://aws.amazon.com/cli/.

- > **Note**
  >
  > Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr create-cluster --instance-type m5.xlarge --instance-count 3 --emrfs
 Consistent=true \
--release-label emr-7.9.0 --ec2-attributes KeyName=myKey
```

Para verificar se a visualização consistente está ativada usando o AWS Management Console

- Para verificar se a exibição consistente está habilitada no console, navegue até Cluster List (Lista de clusters) e selecione o nome do cluster para visualizar Cluster Details (Detalhes do cluster). O campo "EMRFS consistent view" tem um valor de `Enabled` ou `Disabled`.

Para verificar se a exibição consistente está habilitada, examinando o arquivo **`emrfs-site.xml`**

- Você pode verificar se a consistência está habilitada inspecionando o arquivo de configuração `emrfs-site.xml` no nó principal do cluster. Se o valor booleano para `fs.s3.consistent`

estiver definido como `true`, a visualização consistente estará habilitada para operações do sistema de arquivo envolvendo o Amazon S3.

Noções básicas sobre como a visualização consistente do EMRFS rastreia objetos no Amazon S3

O EMRFS cria uma visualização consistente de objetos no Amazon S3, adicionando informações sobre esses objetos aos metadados do EMRFS. O EMRFS adiciona essas listagens para seus metadados quando:

- Um objeto gravado pelo EMRFS no decurso de um trabalho do Amazon EMR.
- Um objeto é sincronizado com ou importado para os metadados do EMRFS usando a CLI do EMRFS.

Objetos lidos pelo EMRFS não são automaticamente adicionados aos metadados. Quando o EMRFS exclui um objeto, ainda permanece uma listagem nos metadados com um estado excluído até que essa listagem seja limpa usando a CLI do EMRFS. Para saber mais sobre a CLI, consulte Referência de comando da CLI do EMRFS. Para obter mais informações sobre como limpar listas nos metadados do EMRFS, consulte Metadados da visualização consistente do EMRFS.

Para cada operação do Amazon S3, o EMRFS verifica os metadados para obter informações sobre o conjunto de objetos na visualização consistente. Se o EMRFS descobrir que o Amazon S3 está inconsistente durante uma dessas operações, ele repetirá a operação de acordo com os parâmetros definidos nas propriedades da configuração de `emrfs-site`. Depois de esgotar as tentativas, o EMRFS lançará uma `ConsistencyException` ou registrará a exceção em log e continuará o fluxo de trabalho. Para obter mais informações sobre a lógica de novas tentativas, consulte Lógica de novas tentativas. Você pode encontrar `ConsistencyExceptions` nos seus logs, por exemplo:

- listStatus: nenhum objeto do Amazon S3 para o item de metadados `/S3_bucket/dir/object`
- getFileStatus: a chave `dir/file` está presente nos metadados, mas não no Amazon S3

Se você excluir um objeto diretamente do Amazon S3 que é acompanhado pela visualização de consistência do EMRFS, o EMRFS tratará o objeto como inconsistente, pois ele ainda estará listado nos metadados como presente no Amazon S3. Se os metadados ficarem fora de sincronia com os objetos que estão sendo acompanhados pelo EMRFS no Amazon S3, você poderá usar o subcomando sync da CLI do EMRFS para redefinir os metadados para que reflitam o Amazon S3.

Para descobrir discrepâncias entre os metadados e o Amazon S3, use o diff. Por fim, o EMRFS só tem uma visualização consistente dos objetos referenciados nos metadados. Pode haver outros objetos no mesmo caminho do Amazon S3 que não estão sendo controlados. Quando o EMRFS lista os objetos em um caminho do Amazon S3, ele retorna o superconjunto dos objetos que estão sendo controlados nos metadados e aqueles no caminho do Amazon S3.

Lógica de novas tentativas

O EMRFS tenta verificar a consistência das listas para objetos controlados em seus metadados por um número específico de novas tentativas. O padrão é 5. No caso em que o número de novas tentativas for excedido, o trabalho originário retornará uma falha, a não ser que `fs.s3.consistent.throwExceptionOnInconsistency` seja definido como `false`, onde ele registrará somente os objetos controlados como inconsistentes. O EMRFS usa uma política de nova tentativa de recuo exponencial por padrão, mas você também pode configurá-lo para uma política fixa. Os usuários também podem querer fazer novas tentativas durante um determinado período antes de prosseguirem com o restante do seu trabalho sem lançar uma exceção. Eles podem conseguir isso definindo `fs.s3.consistent.throwExceptionOnInconsistency` como `false`, `fs.s3.consistent.retryPolicyType` como `fixed` e `fs.s3.consistent.retryPeriodSeconds` para o valor desejado. O exemplo a seguir cria um cluster com a consistência habilitada, o que registra as inconsistências e define um intervalo de repetição fixo de 10 segundos:

Example Definir o período de repetição como um valor fixo

```
aws emr create-cluster --release-label emr-7.9.0 \
--instance-type m5.xlarge --instance-count 1 \
--emrfs Consistent=true,Args=[fs.s3.consistent.throwExceptionOnInconsistency=false,
 fs.s3.consistent.retryPolicyType=fixed,fs.s3.consistent.retryPeriodSeconds=10] --ec2-
attributes KeyName=myKey
```

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

Para obter mais informações, consulte Consistent view.

As configurações do EMRFS para IMDS recebem chamadas regionais

O EMRFS conta com o IMDS (serviço de metadados de instância) para obter a região da instância e endpoints do Amazon S3, do DynamoDB ou do. AWS KMS No entanto, o IMDS tem um limite de quantas solicitações poderá processar, e as solicitações que excederem esse limite falharão. Esse limite de IMDS pode causar falhas na inicialização do EMRFS e fazer a consulta ou o comando falhar. É possível usar o mecanismo de repetição de recuo exponencial aleatório a seguir e as propriedades de configuração de uma região de fallback em emrfs-site.xml para resolver o cenário em que todas as tentativas falham.

```
<property>
    <name>fs.s3.region.retryCount</name>
    <value>3</value>
    <description>
    Maximum retries that would be attempted to get AWS region.
    </description>
</property>
<property>
    <name>fs.s3.region.retryPeriodSeconds</name>
    <value>3</value>
    <description>
    Base sleep time in second for each get-region retry.
    </description>
</property>
<property>
    <name>fs.s3.region.fallback</name>
    <value>us-east-1</value>
    <description>
    Fallback to this region after maximum retries for getting AWS region have been
 reached.
    </description>
</property>
```

Metadados da visualização consistente do EMRFS

A visualização consistente do EMRFS acompanha a consistência usando uma tabela do DynamoDB para acompanhar objetos no Amazon S3 que foram sincronizados com o EMRFS ou criados pelo EMRFS. Os metadados são usados para controlar todas as operações (ler, gravar, atualizar e copiar), e nenhum conteúdo real é armazenado neles. Esses metadados são usados para validar se os objetos ou metadados recebidos do Amazon S3 correspondem ao que se espera. Essa confirmação dá ao EMRFS a capacidade de verificar a consistência e read-after-write a consistência

de listas para novos objetos que o EMRFS grava no Amazon S3 ou para objetos sincronizados com o EMRFS. Vários clusters podem compartilhar os mesmos metadados.

Como adicionar entradas aos metadados

Você pode usar os subcomandos `sync` ou `import` para adicionar entradas aos metadados. `sync` reflete o estado dos objetos do Amazon S3 em um caminho, enquanto `import` é usado estritamente para adicionar novas entradas aos metadados. Para obter mais informações, consulte Referência de comando da CLI do EMRFS.

Como verificar diferenças entre metadados e objetos no Amazon S3

Para verificar se há diferenças entre os metadados e o Amazon S3, use o subcomando `diff` da CLI do EMRFS. Para obter mais informações, consulte Referência de comando da CLI do EMRFS.

Como saber se as operações de metadados estão sendo limitadas

O EMRFS define limites de capacidade de throughput padrão nos metadados para suas operações de leitura e gravação em 500 e 100 unidades, respectivamente. Grandes número de objetos ou buckets podem fazer com que as operações excedam essa capacidade e, nesse ponto, o DynamoDB limitará as operações. Por exemplo, uma aplicação pode fazer com que o EMRFS lance um `ProvisionedThroughputExceededException` se você executar uma operação que exceda os limites de capacidade. Em caso de controle de utilização, a ferramenta de CLI do EMRFS tenta repetir a gravação na tabela do DynamoDB usando o recuo exponencial até que a operação seja concluída ou quando atingir o valor máximo de repetição para gravar objetos do Amazon EMR no Amazon S3.

É possível configurar seus próprios limites de capacidade de throughput. No entanto, o DynamoDB tem limites de partição estritos de três mil unidades de capacidade de leitura RCUs () e mil unidades de capacidade de gravação WCUs () por segundo para operações de leitura e gravação. Para evitar `sync` falhas causadas por controle de utilização, recomenda-se limitar o throughput das operações de leitura a menos de três mil RCUs e as operações de gravação a menos de mil. WCUs Para obter instruções sobre como definir limites personalizados de capacidade de throughput, consulte Configurar a visualização consistente.

Você também pode visualizar CloudWatch métricas da Amazon para os seus metadados do EMRFS no console do DynamoDB, onde é possível ver o número de solicitações de leitura e gravação limitadas. Se você tiver um valor diferente de zero para solicitações limitadas, seu aplicativo poderá se beneficiar com o aumento da capacidade de throughput alocada para operações de leitura ou gravação. Você também poderá perceber um benefício de desempenho se constatar que as suas

operações estão se aproximando da capacidade de throughput máxima alocada em leituras ou gravações por um período prolongado.

Características de throughput para operações notáveis do EMRFS

O padrão para operações de leitura e gravação é de 400 e 100 unidades de capacidade de throughput, respectivamente. As seguintes características de desempenho dão uma ideia de qual throughput é necessário para determinadas operações. Esses testes foram realizados usando um cluster `m3.large` de nó único. Todas as operações foram definidas em um único thread. O desempenho tem uma variação considerável com base em características específicas dos aplicativos e pode exigir experimentação para otimizar as operações do sistema de arquivos.

Operação	Média read-per-second	Média write-per-second
create (objeto)	26.79	6.70
delete (objeto)	10.79	10.79
delete (diretório contendo 1000 objetos)	21.79	338.40
getFileStatus(objeto)	34.70	0
getFileStatus(diretório)	19.96	0
listStatus (diretório contendo 1 objeto)	43.31	0
listStatus (diretório contendo 10 objetos)	44.34	0
listStatus (diretório contendo 100 objetos)	84.44	0
listStatus (diretório contendo 1.000 objetos)	308.81	0
listStatus (diretório contendo 10.000 objetos)	416.05	0

Operação	Média read-per-second	Média write-per-second
listStatus (diretório contendo 100.000 objetos)	823.56	0
listStatus (diretório contendo 1M de objetos)	882.36	0
mkdir (contínuo por 120 segundos)	24.18	4.03
mkdir	12.59	0
rename (objeto)	19.53	4.88
rename (diretório contendo 1.000 objetos)	23.22	339.34

Para enviar uma etapa que limpa dados antigos do seu repositório de metadados

Os usuários podem querer remover entradas específicas nos metadados baseados no DynamoDB. Isso pode ajudar a reduzir os custos de armazenamento associados à tabela. Os usuários têm a capacidade de limpar entradas particulares manualmente ou programaticamente, usando o subcomando `delete` da CLI do EMRFS. No entanto, se você excluir entradas dos metadados, o EMRFS deixará de fazer verificações de consistência.

A limpeza usando um programa após a conclusão de um trabalho pode ser feita mediante o envio de uma etapa final ao cluster, que executa um comando na CLI do EMRFS. Por exemplo, digite o seguinte comando para enviar uma etapa ao seu cluster para excluir todas as entradas com mais de dois dias.

```
aws emr add-steps --cluster-id j-2AL4XXXXXX5T9 --steps Name="emrfsCLI",Jar="command-
runner.jar",Args=["emrfs","delete","--time","2","--time-unit","days"]
{
    "StepIds": [
        "s-B12345678902"
    ]
}
```

Use o StepId valor retornado para verificar os logs em busca do resultado da operação.

Configurar notificações de consistência para CloudWatch o Amazon SQS

Você pode habilitar CloudWatch métricas e mensagens do Amazon SQS no EMRFS para problemas de consistência eventual do Amazon S3.

CloudWatch

Quando CloudWatch as métricas estão habilitadas, uma métrica denominada Inconsistency é enviada sempre que uma chamada de `FileSystem` API falha devido à consistência eventual do Amazon S3.

Visualizar as CloudWatch métricas de problemas de consistência eventual do Amazon S3

Para visualizar a métrica de inconsistência no CloudWatch console, selecione as métricas do EMRFS e, em seguida, selecione um par JobFlowId/Nome da métrica. Por exemplo: `j-162XXXXXXM2CU ListStatus`, `j-162XXXXXXM2CU GetFileStatus` e assim por diante.

1. Abra o CloudWatch console em https://console.aws.amazon.com/cloudwatch/.
2. Em Dashboard (Painel), na seção Metrics (Métricas), escolha EMRFS.
3. No painel Job Flow Metrics, selecione um ou mais pares JobFlowId/Metric Name. Uma representação gráfica das métricas aparece na janela abaixo.

Amazon SQS

Quando as notificações do Amazon SQS estiverem habilitadas, será criada uma fila do Amazon SQS chamada `EMRFS-Inconsistency-<jobFlowId>` quando o EMRFS for inicializado. As mensagens do Amazon SQS são enviadas à fila quando uma chamada de API `FileSystem` falha devido à consistência eventual do Amazon S3. A mensagem contém informações como API JobFlowId, uma lista de caminhos inconsistentes, um rastreamento de pilha e assim por diante. As mensagens podem ser lidas usando o console do Amazon SQS ou o comando `read-sqs` do EMRFS.

Gerenciar mensagens do Amazon SQS para possíveis problemas de consistência do Amazon S3

As mensagens do Amazon SQS para problemas de consistência eventual do Amazon S3 podem ser lidas usando a CLI do EMRFS. Para ler mensagens de uma fila do Amazon SQS do EMRFS, digite o comando `read-sqs` e especifique um local de saída no sistema de arquivos local do nó principal para o arquivo de saída resultante.

Você também pode excluir uma fila do Amazon SQS do EMRFS usando o comando `delete-sqs`.

1. Para ler mensagens de uma fila do Amazon SQS, digite o comando a seguir. *queuename*Substitua pelo nome da fila do Amazon SQS que você configurou e */path/filename* substitua pelo caminho do arquivo de saída:

   ```
   emrfs read-sqs --queue-name queuename --output-file /path/filename
   ```

   Por exemplo, para ler e gerar a saída de mensagens do Amazon SQS na fila padrão, digite:

   ```
   emrfs read-sqs --queue-name EMRFS-Inconsistency-j-162XXXXXXM2CU --output-file /path/filename
   ```

   > **Note**
   >
   > Você também pode usar os atalhos `-q` e `-o` em vez de `--queue-name` e `--output-file`, respectivamente.

2. Para excluir uma fila do Amazon SQS digite o seguinte comando:

   ```
   emrfs delete-sqs --queue-name queuename
   ```

   Por exemplo, para excluir a fila padrão, digite:

   ```
   emrfs delete-sqs --queue-name EMRFS-Inconsistency-j-162XXXXXXM2CU
   ```

   > **Note**
   >
   > Você também pode usar o atalho `-q` no lugar de `--queue-name`.

Configurar a visualização consistente

Você pode definir configurações adicionais para a visualização consistente usando propriedades de configuração para propriedades do `emrfs-site`. Por exemplo, você pode escolher um throughput do DynamoDB diferente fornecendo os seguintes argumentos para a opção `--emrfs` da CLI usando

a classificação da configuração do emrfs-site (somente Amazon EMR versão 4.x ou posterior) ou uma ação de bootstrap para configurar o arquivo emrfs-site.xml no nó principal:

Example Alterar os valores de leitura e gravação de metadados padrão na execução do cluster

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge \
--emrfs Consistent=true,Args=[fs.s3.consistent.metadata.read.capacity=600,\
fs.s3.consistent.metadata.write.capacity=300] --ec2-attributes KeyName=myKey
```

Como alternativa, use o seguinte arquivo de configuração e salve-o localmente ou no Amazon S3:

```
[
    {
      "Classification": "emrfs-site",
      "Properties": {
        "fs.s3.consistent.metadata.read.capacity": "600",
        "fs.s3.consistent.metadata.write.capacity": "300"
      }
    }
 ]
```

Use a configuração que você criou com a seguinte sintaxe:

```
aws emr create-cluster --release-label emr-7.9.0 --applications Name=Hive \
--instance-type m5.xlarge --instance-count 2 --configurations file://./myConfig.json
```

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

As opções a seguir podem ser definidas usando configurações ou AWS CLI `--emrfs` argumentos. Para obter mais informações sobre esses argumentos, consulte AWS CLI Command Reference.

Propriedades **`emrfs-site.xml`** para a visualização consistente

Propriedade	Valor padrão	Descrição
`fs.s3.consistent`	**`false`**	

Propriedade	Valor padrão	Descrição
		Quando definida como **`true`**, essa propriedade configura o EMRFS usar o DynamoDB para fornecer consistência.
`fs.s3.consistent.retryPolicyType`	**`exponential`**	Essa propriedade identifica a política a ser usada em caso de novas tentativas para problemas de consistência. As opções incluem: exponential, fixed ou none.
`fs.s3.consistent.retryPeriodSeconds`	**`1`**	Essa propriedade define o tempo de espera entre tentativas de repetição de consistência.
`fs.s3.consistent.retryCount`	**`10`**	Essa propriedade define o número máximo de tentativas quando é detectada uma inconsistência.
`fs.s3.consistent.throwExceptionOnInconsistency`	**`true`**	Essa propriedade determina se uma exceção de consistência deve ser lançada ou registrada. Quando definida como **`true`**, um `ConsistencyException` é lançado.
`fs.s3.consistent.metadata.autoCreate`	**`true`**	Quando definida como **`true`**, essa propriedade permite a criação automática de tabelas de metadados.

Propriedade	Valor padrão	Descrição
`fs.s3.consistent.metadata.etag.verification.enabled`	**`true`**	Com o Amazon EMR 5.29.0, essa propriedade é habilitada por padrão. Quando habilitado, o EMRFS usa o S3 ETags para verificar se os objetos que estão sendo lidos estão na versão mais recente disponível. Esse atributo é útil para casos de read-after-update uso em que arquivos no S3 estão sendo substituídos, mas mantêm o mesmo nome. Atualmente, esse recurso de ETag verificação não funciona com o S3 Select.
`fs.s3.consistent.metadata.tableName`	**`EmrFSMetadata`**	Essa propriedade especifica o nome da tabela de metadados no DynamoDB.
`fs.s3.consistent.metadata.read.capacity`	**`500`**	Essa propriedade especific a a capacidade de leitura do DynamoDB a ser provisionada quando a tabela de metadados é criada.
`fs.s3.consistent.metadata.write.capacity`	**`100`**	Essa propriedade especifica a capacidade de gravação do DynamoDB a ser provisionada quando a tabela de metadados é criada.

Propriedade	Valor padrão	Descrição
`fs.s3.consistent.fastList`	**`true`**	Quando definida como **`true`**, essa propriedade usa vários threads para listar um diretório (quando necessário). A consistência deve ser habilitada para usar essa propriedade.
`fs.s3.consistent.fastList.prefetchMetadata`	**`false`**	Quando definida como **`true`**, essa propriedade permite que metadados façam uma pré-busca por diretórios que contêm mais de 20.000 itens.
`fs.s3.consistent.notification.CloudWatch`	**`false`**	Quando definida como**`true`**, CloudWatch as métricas são habilitadas para chamadas de FileSystem API que falham devido a problemas de consistência eventual do Amazon S3.
`fs.s3.consistent.notification.SQS`	**`false`**	Quando definida como **`true`**, notificações de consistência eventual são enviadas a uma fila do Amazon SQS.
`fs.s3.consistent.notification.SQS.queueName`	**`EMRFS-Inconsistency-<jobFlowId>`**	Alterar essa propriedade permite que você especifique seu próprio nome de fila do SQS para mensagens a respeito de problemas de consistência eventual do Amazon S3.

Propriedade	Valor padrão	Descrição
`fs.s3.consistent.notification.SQS.customMsg`	**`none`**	Essa propriedade permite especificar informações personalizadas incluídas em mensagens do SQS a respeito de problemas de consistência eventual do Amazon S3. Se um valor não for especificado para essa propriedade, o campo correspondente na mensagem ficará vazio.
`fs.s3.consistent.dynamodb.endpoint`	**`none`**	Essa propriedade permite especificar um endpoint personalizado do DynamoDB para as seus metadados de exibição consistente.
`fs.s3.useRequesterPaysHeader`	**`false`**	Quando definida como **`true`**, essa propriedade permite solicitações do Amazon S3 para buckets com a opção de pagante de solicitações habilitada.

Referência de comando da CLI do EMRFS

A CLI do EMRFS é instalada por padrão em todos os nós principais do cluster criados usando o Amazon EMR versão 3.2.1 ou posterior. Você pode usar a CLI do EMRFS para gerenciar os metadados para visualização consistente.

> **Note**
>
> O emrfs comando só é suportado com a emulação do terminal VT1 00. No entanto, ele pode funcionar com outros modos de emulador de terminal.

comando de nível superior emrfs

O comando de nível superior emrfs oferece suporte à seguinte estrutura.

```
emrfs [describe-metadata | set-metadata-capacity | delete-metadata | create-metadata |
 \
list-metadata-stores | diff | delete | sync | import ] [options] [arguments]
```

Especifique [opções], com ou sem [argumentos] conforme descrito na tabela a seguir. Para [opções] específicas de subcomandos (`describe-metadata`, `set-metadata-capacity`, etc.), consulte cada subcomando abaixo.

[Opções] para emrfs

Opção	Descrição	Obrigatório
`-a` *AWS_ACCESS_KEY_ID* | `--access-key` *AWS_ACCESS_KEY_ID*	A chave de AWS acesso da que você usa para gravar objetos no Amazon S3 e para criar ou acessar um repositório de metadados no DynamoDB. Por padrão, *AWS_ACCESS_KEY_ID* é definido como a chave de acesso usada para criar o cluster.	Não
`-s` *AWS_SECRET_ACCESS_KEY* | `--secret-key` *AWS_SECRET_ACCESS_KEY*	A chave AWS secreta da associada à chave de acesso que você usa para gravar objetos no Amazon S3 e para criar ou acessar um repositório de metadados no DynamoDB. Por padrão, *AWS_SECRET_ACCESS_KEY* é definida como a chave secreta associada à chave de acesso usada para criar o cluster.	Não
`-v` | `--verbose`	Torna a saída detalhada.	Não
`-h` | `--help`	Exibe a mensagem de ajuda para o comando `emrfs` com uma instrução de uso.	Não

subcomando emrfs describe-metadata

[Opções] para emrfs describe-metadata

Opção	Descrição	Obrigatório
`-m` *METADATA_NAME* | `--metadata-name` *METADATA_NAME*	*METADATA_NAME* é o nome da tabela de metadados do DynamoDB. Se o *METADATA_NAME* argumento não for fornecido, o valor padrão será`EmrFSMetadata` .	Não

Example exemplo de emrfs describe-metadata

O exemplo a seguir descreve a tabela de metadados padrão.

```
$ emrfs describe-metadata
EmrFSMetadata
  read-capacity: 400
  write-capacity: 100
  status: ACTIVE
  approximate-item-count (6 hour delay): 12
```

subcomando emrfs set-metadata-capacity

[Opções] para emrfs set-metadata-capacity

Opção	Descrição	Obrigatório
`-m` *METADATA_NAME* | `--metadata-name` *METADATA_NAME*	*METADATA_NAME* é o nome da tabela de metadados do DynamoDB. Se o *METADATA_NAME* argumento não for fornecido, o valor padrão será`EmrFSMetadata` .	Não
	A capacidade de throughput de leitura solicitada para a tabela de metadados. Se o *READ_CAPA*	Não

Opção	Descrição	Obrigatório
`-r READ_CAPACITY | --read-capacity READ_CAPACITY`	*CITY* argumento não for fornecido, o valor padrão será`400`.	
`-w WRITE_CAPACITY | --write-capacity WRITE_CAPACITY`	A capacidade de throughput de gravação solicitada para a tabela de metadados. Se o *WRITE_CAPACITY* argumento não for fornecido, o valor padrão será`100`.	Não

Example exemplo de emrfs set-metadata-capacity

O exemplo a seguir define a capacidade de throughput de leitura como `600` e a capacidade de gravação como `150` para uma tabela de metadados denominada `EmrMetadataAlt`.

```
$ emrfs set-metadata-capacity --metadata-name EmrMetadataAlt  --read-capacity 600 --
write-capacity 150
  read-capacity: 400
  write-capacity: 100
  status: UPDATING
  approximate-item-count (6 hour delay): 0
```

subcomando emrfs delete-metadata

[Opções] para emrfs delete-metadata

Opção	Descrição	Obrigatório
`-m METADATA_NAME | --metadata-name METADATA_NAME`	*METADATA_NAME* é o nome da tabela de metadados do DynamoDB. Se o *METADATA_NAME* argumento não for fornecido, o valor padrão seráEmrFSMetadata .	Não

Example exemplo de delete-metadata

O exemplo a seguir exclui a tabela de metadados padrão.

```
$ emrfs delete-metadata
```

subcomando emrfs create-metadata

[Opções] para emrfs create-metadata

Opção	Descrição	Obrigatório
`-m` *METADATA_NAME* | `--metadata-name` *METADATA_NAME*	*METADATA_NAME* é o nome da tabela de metadados do DynamoDB. Se o *METADATA_NAME* argumento não for fornecido, o valor padrão seráEmrFSMetadata .	Não
`-r` *READ_CAPACITY* | `--read-capacity` *READ_CAPACITY*	A capacidade de throughput de leitura solicitada para a tabela de metadados. Se o *READ_CAPACITY* argumento não for fornecido, o valor padrão será400.	Não
`-w` *WRITE_CAPACITY* | `--write-capacity` *WRITE_CAPACITY*	A capacidade de throughput de gravação solicitada para a tabela de metadados. Se o *WRITE_CAPACITY* argumento não for fornecido, o valor padrão será100.	Não

Example exemplo de emrfs create-metadata

O exemplo a seguir cria uma tabela de metadados chamada de EmrFSMetadataAlt.

```
$ emrfs create-metadata -m EmrFSMetadataAlt
Creating metadata: EmrFSMetadataAlt
EmrFSMetadataAlt
  read-capacity: 400
  write-capacity: 100
```

```
  status: ACTIVE
  approximate-item-count (6 hour delay): 0
```

subcomando emrfs list-metadata-stores

O subcomando emrfs list-metadata-stores não tem [opções].

Example List-metadata-stores exemplo

O exemplo a seguir lista suas tabelas de metadados.

```
$ emrfs list-metadata-stores
  EmrFSMetadata
```

subcomando emrfs diff

[Opções] para emrfs diff

Opção	Descrição	Obrigatório
`-m` *METADATA_NAME* | `--metadata-name` *METADATA_NAME*	*METADATA_NAME* é o nome da tabela de metadados do DynamoDB. Se o *METADATA_NAME* argumento não for fornecido, o valor padrão será`EmrFSMetadata` .	Não
s3://s3Path	O caminho para o bucket do Amazon S3, para a comparação com a tabela de metadados. Buckets são sincronizados recursivamente.	Sim

Example exemplo de emrfs diff

O exemplo a seguir compara a tabela de metadados padrão com um bucket do Amazon S3.

```
$ emrfs diff s3://elasticmapreduce/samples/cloudfront
BOTH | MANIFEST ONLY | S3 ONLY
DIR elasticmapreduce/samples/cloudfront
DIR elasticmapreduce/samples/cloudfront/code/
```

```
DIR elasticmapreduce/samples/cloudfront/input/
DIR elasticmapreduce/samples/cloudfront/logprocessor.jar
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-14.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-15.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-16.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-17.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-18.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-19.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-20.WxYz1234
DIR elasticmapreduce/samples/cloudfront/code/cloudfront-loganalyzer.tgz
```

subcomando emrfs delete

[Opções] para emrfs delete

Opção	Descrição	Obrigatório
`-m` *METADATA_NAME* | `--metadata-name` *METADATA_NAME*	*METADATA_NAME* é o nome da tabela de metadados do DynamoDB. Se o *METADATA_NAME* argumento não for fornecido, o valor padrão será`EmrFSMetadata` .	Não
s3://s3Path	O caminho para o bucket do Amazon S3 que você está rastreando para uma visualização consistente. Buckets são sincronizados recursivamente.	Sim
`-t` *TIME* | `--time` *TIME*	O tempo de expiração (interpretado usando o argumento de unidade de tempo). Todas as entradas de metadados anteriores ao *TIME* argumento são excluídas do bucket especificado.	
`-u` *UNIT* | `--time-unit` *UNIT*	A medida usada para interpretar o argumento TIME (nanossegundos, microssegundos, milissegundos, segundos, minutos, horas ou dias). Se nenhum argumento for especificado, o valor padrão será `days`.	

Opção	Descrição	Obrigatório
`--read-consumption` *READ_CONSUMPTION*	A quantidade solicitada de throughput de leitura disponível utilizado para a operação delete. Se o *READ_CONSUMPTION* argumento não for especificado, o valor padrão será `400`.	Não
`--write-consumption` *WRITE_CONSUMPTION*	A quantidade solicitada de throughput de gravação disponível utilizado para a operação delete. Se o *WRITE_CONSUMPTION* argumento não for especificado, o valor padrão será `100`.	Não

Example exemplo de emrfs delete

O exemplo a seguir remove todos os objetos em um bucket do Amazon S3 dos metadados de rastreamento para uma visualização consistente.

```
$ emrfs delete s3://elasticmapreduce/samples/cloudfront
entries deleted: 11
```

subcomando emrfs import

[Opções] para emrfs import

Opção	Descrição	Obrigatório
`-m` *METADATA_NAME* | `--metadata-name` *METADATA_NAME*	*METADATA_NAME* é o nome da tabela de metadados do DynamoDB. Se o *METADATA_NAME* argumento não for fornecido, o valor padrão será`EmrFSMetadata` .	Não
s3://s3Path		Sim

Opção	Descrição	Obrigatório
	O caminho para o bucket do Amazon S3 que você está rastreando para uma visualização consistente. Buckets são sincronizados recursivamente.	
`--read-consumption` *READ_CONSUMPTION*	A quantidade solicitada de throughput de leitura disponível utilizado para a operação delete. Se o *READ_CONSUMPTION* argumento não for especificado, o valor padrão será `400`.	Não
`--write-consumption` *WRITE_CONSUMPTION*	A quantidade solicitada de throughput de gravação disponível utilizado para a operação delete. Se o *WRITE_CONSUMPTION* argumento não for especificado, o valor padrão será `100`.	Não

Example exemplo de emrfs import

O exemplo a seguir importa todos os objetos em um bucket do Amazon S3 com os metadados de rastreamento para uma visualização consistente. Todas as chaves desconhecidas são ignoradas.

```
$ emrfs import s3://elasticmapreduce/samples/cloudfront
```

subcomando emrfs sync

[Opções] para emrfs sync

Opção	Descrição	Obrigatório
`-m` *METADATA_NAME* | `--metadata-name` *METADATA_NAME*	*METADATA_NAME* é o nome da tabela de metadados do DynamoDB. Se o *METADATA_NAME* argumento não for fornecido, o valor padrão será`EmrFSMetadata` .	Não

Opção	Descrição	Obrigatório
s3://s3Path	O caminho para o bucket do Amazon S3 que você está rastreando para uma visualização consistente. Buckets são sincronizados recursivamente.	Sim
`--read-consumption` *READ_CONSUMPTION*	A quantidade solicitada de throughput de leitura disponível utilizado para a operação delete. Se o *READ_CONSUMPTION* argumento não for especificado, o valor padrão será `400`.	Não
`--write-consumption` *WRITE_CONSUMPTION*	A quantidade solicitada de throughput de gravação disponível utilizado para a operação delete. Se o *WRITE_CONSUMPTION* argumento não for especificado, o valor padrão será `100`.	Não

Example exemplo do comando emrfs sync

O exemplo a seguir importa todos os objetos em um bucket do Amazon S3 com os metadados de rastreamento para uma visualização consistente. Todas as chaves desconhecidas são excluídas.

```
$ emrfs sync s3://elasticmapreduce/samples/cloudfront
Synching samples/cloudfront                                    0 added | 0 updated |
 0 removed | 0 unchanged
Synching samples/cloudfront/code/                              1 added | 0 updated |
 0 removed | 0 unchanged
Synching samples/cloudfront/                                   2 added | 0 updated |
 0 removed | 0 unchanged
Synching samples/cloudfront/input/                             9 added | 0 updated |
 0 removed | 0 unchanged
Done synching s3://elasticmapreduce/samples/cloudfront         9 added | 0 updated |
 1 removed | 0 unchanged
creating 3 folder key(s)
folders written: 3
```

subcomando emrfs read-sqs

[Opções] para emrfs read-sqs

Opção	Descrição	Obrigatório
`-q` *QUEUE_NAME* | `--queue-name` *QUEUE_NAME*	*QUEUE_NAME* é o nome da fila do Amazon SQS configurada em. `emrfs-site.xml` O valor padrão é **EMRFS-Inconsistency-<jobFlowId>** .	Sim
`-o` *OUTPUT_FILE* | `--output-file` *OUTPUT_FILE*	*OUTPUT_FILE* é o caminho para o arquivo de saída no sistema de arquivos local do nó principal . Mensagens lidas na fila são gravadas nesse arquivo.	Sim

subcomando emrfs delete-sqs

[Opções] para emrfs delete-sqs

Opção	Descrição	Obrigatório
`-q` *QUEUE_NAME* | `--queue-name` *QUEUE_NAME*	*QUEUE_NAME* é o nome da fila do Amazon SQS configurada em. `emrfs-site.xml` O valor padrão é **EMRFS-Inconsistency-<jobFlowId>** .	Sim

Enviar comandos da CLI do EMRFS como etapas

O exemplo a seguir mostra como usar o `emrfs` utilitário no nó principal aproveitando a API AWS CLI or e a `command-runner.jar` para executar o `emrfs` comando como uma etapa. O exemplo usa o AWS SDK para Python (Boto3) para adicionar uma etapa a um cluster que adiciona objetos em um bucket do Amazon S3 para a tabela de metadados padrão do EMRFS.

```
import boto3
from botocore.exceptions import ClientError


def add_emrfs_step(command, bucket_url, cluster_id, emr_client):
    """
    Add an EMRFS command as a job flow step to an existing cluster.

    :param command: The EMRFS command to run.
    :param bucket_url: The URL of a bucket that contains tracking metadata.
    :param cluster_id: The ID of the cluster to update.
    :param emr_client: The Boto3 Amazon EMR client object.
    :return: The ID of the added job flow step. Status can be tracked by calling
             the emr_client.describe_step() function.
    """
    job_flow_step = {
        "Name": "Example EMRFS Command Step",
        "ActionOnFailure": "CONTINUE",
        "HadoopJarStep": {
            "Jar": "command-runner.jar",
            "Args": ["/usr/bin/emrfs", command, bucket_url],
        },
    }

    try:
        response = emr_client.add_job_flow_steps(
            JobFlowId=cluster_id, Steps=[job_flow_step]
        )
        step_id = response["StepIds"][0]
        print(f"Added step {step_id} to cluster {cluster_id}.")
    except ClientError:
        print(f"Couldn't add a step to cluster {cluster_id}.")
        raise
    else:
        return step_id


def usage_demo():
    emr_client = boto3.client("emr")
    # Assumes the first waiting cluster has EMRFS enabled and has created metadata
    # with the default name of 'EmrFSMetadata'.
    cluster = emr_client.list_clusters(ClusterStates=["WAITING"])["Clusters"][0]
    add_emrfs_step(
```

```
        "sync", "s3://elasticmapreduce/samples/cloudfront", cluster["Id"], emr_client
    )


if __name__ == "__main__":
    usage_demo()
```

Você pode usar o valor de `step_id` retornado para verificar os logs em busca do resultado da operação.

Autorizar acesso aos dados do EMRFS no Amazon S3

Por padrão, o perfil do EMR EC2 determina as permissões para acesso aos dados do EMRFS no Amazon S3. As políticas do IAM que são anexadas a esse perfil se aplicam independentemente do usuário ou do grupo que faz a solicitação por meio do EMRFS. O padrão é `EMR_EC2_DefaultRole`. Para obter mais informações, consulte Função de serviço para EC2 instâncias de cluster (perfil de EC2 instância).

A partir do Amazon EMR versão 5.10.0, você pode usar uma configuração de segurança para especificar perfis do IAM para o EMRFS. Com isso, você pode personalizar permissões para solicitações do EMRFS ao Amazon S3 para clusters com vários usuários. Você pode especificar perfis do IAM diferentes para usuários e grupos diferentes, e para diferentes locais de bucket do Amazon S3 com base no prefixo no Amazon S3. Quando o EMRFS faz uma solicitação ao Amazon S3 que corresponde a usuários, grupos ou locais especificados, o cluster usa a perfil correspondente que você especifica em vez do perfil do EMR. EC2 Para obter mais informações, consulte Configure IAM roles for EMRFS requests to Amazon S3.

Como alternativa, se sua solução do Amazon EMR solicitar algo além do que é fornecido pelos perfis do IAM para EMRFS, você poderá definir uma classe de provedor de credenciais personalizado, que permita personalizar o acesso aos dados do EMRFS no Amazon S3.

Criar um provedor de credenciais personalizado para dados do EMRFS no Amazon S3

Para criar um provedor de credenciais personalizado, você implementa as classes AWSCredentialsProvider e Hadoop Configurable.

Para obter uma explicação detalhada dessa abordagem, consulte Securely analyze data from another AWS account with EMRFS no blog de big data da AWS . A postagem do blog inclui um

tutorial que mostra o processoend-to-end, desde a criação de perfis do IAM até a execução do cluster. Ela também fornece um exemplo de código Java que implementa a classe de provedor de credenciais personalizado.

As etapas básicas são:

Para especificar um provedor de credenciais personalizado

1. Crie uma classe de provedor de credenciais personalizado compilada como um arquivo JAR.
2. Execute um script como uma ação de bootstrap para copiar o arquivo JAR do provedor de credenciais personalizado no local `/usr/share/aws/emr/emrfs/auxlib` no nó principal do cluster. Para obter mais informações sobre ações de bootstrap, consulte (Optional) Create bootstrap actions to install additional software.
3. Personalize a classificação `emrfs-site` para especificar a classe que você implementa no arquivo JAR. Para obter mais informações sobre como especificar objetos de configuração para personalizar as aplicações, consulte Configuring applications no Guia de lançamento do Amazon EMR.

   O exemplo a seguir demonstra um comando `create-cluster` que executa um cluster do Hive com parâmetros de configuração comuns e também inclui:

   - Uma ação de bootstrap que executa o script, *`copy_jar_file.sh`*, que é salvo no *`amzn-s3-demo-bucket`* no Amazon S3.
   - Uma classificação `emrfs-site` que especifica um provedor de credenciais personalizado definido no arquivo JAR como *`MyCustomCredentialsProvider`*

   > **Note**
   >
   > Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr create-cluster --applications Name=Hive \
--bootstrap-actions '[{"Path":"s3://amzn-s3-demo-bucket/
copy_jar_file.sh","Name":"Custom action"}]' \
--ec2-attributes '{"KeyName":"MyKeyPair","InstanceProfile":"EMR_EC2_DefaultRole",\
```

```
"SubnetId":"subnet-xxxxxxxx","EmrManagedSlaveSecurityGroup":"sg-xxxxxxxx",\
"EmrManagedMasterSecurityGroup":"sg-xxxxxxxx"}' \
--service-role EMR_DefaultRole_V2 --enable-debugging --release-label emr-7.9.0 \
--log-uri 's3n://amzn-s3-demo-bucket/' --name 'test-awscredentialsprovider-emrfs' \
--instance-type=m5.xlarge --instance-count 3  \
--configurations '[{"Classification":"emrfs-site",\
"Properties":
{"fs.s3.customAWSCredentialsProvider":"MyAWSCredentialsProviderWithUri"},\
"Configurations":[]}]'
```

Gerenciar o AWS Security Token Service endpoint do padrão

O EMRFS usa o AWS Security Token Service (STS) para recuperar credenciais de segurança temporárias a fim de acessar seus recursos da. AWS As versões anteriores do Amazon EMR enviam todas as AWS STS solicitações do para um único endpoint global em. `https://sts.amazonaws.com` Em vez disso, as versões 5.31.0 e posteriores do Amazon EMR fazem solicitações para endpoints regionais do. AWS STS Isso reduz a latência e melhora a validade do token da sessão. Para obter mais informações sobre AWS STS endpoints, consulte Gerenciando AWS STS em uma AWS região no Guia do AWS Identity and Access Management usuário.

Ao usar as versões 5.31.0 e 6.1.0 e posteriores do Amazon EMR, você pode substituir o endpoint do AWS STS padrão. Para isso, é necessário alterar a propriedade `fs.s3.sts.endpoint` da configuração `emrfs-site`.

O AWS CLI exemplo da a seguir define o AWS STS endpoint do padrão que o EMRFS usa como endpoint global.

```
aws emr create-cluster --release-label <emr-5.33.0> --instance-type m5.xlarge \
--emrfs Args=[fs.s3.sts.endpoint=https://sts.amazonaws.com]
```

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

Como alternativa, é possível criar um arquivo de configuração JSON usando o exemplo a seguir e especificá-lo usando o argumento `--configurations` de `emr create-cluster`. Para obter mais informações sobre o uso de `--configurations`,, consulte AWS CLI Command Reference.

```
[
  {
    "classification": "emrfs-site",
    "properties": {
      "fs.s3.sts.endpoint": "https://sts.amazonaws.com"
    }
  }
]
```

Especificar a criptografia do Amazon S3 usando propriedades do EMRFS

> **Important**
>
> A partir do Amazon EMR versão 4.8.0, você pode usar configurações de segurança para aplicar configurações de criptografia mais facilmente e com mais opções. Recomendamos usar configuração de segurança. Para obter informações, consulte Configure data encryption. As instruções do console descritas nesta seção estão disponíveis para versões anteriores à 4.8.0. Se você usar o AWS CLI para configurar a criptografia do Amazon S3 na configuração de cluster e em uma configuração de segurança em versões subsequentes, a configuração de segurança substituirá a configuração de cluster.

Ao criar um cluster, você pode especificar a criptografia do lado do servidor (SSE) ou a criptografia do lado do cliente (CSE) para dados do EMRFS no emrfs-site usando o console ou as propriedades de classificação por meio da ou do SDK do EMR EMR. `emrfs-site` AWS CLI A SSE e a CSE do Amazon S3 são mutuamente exclusivas. Você pode escolher uma delas, mas não ambas.

Para AWS CLI obter instruções, consulte a seção apropriada para seu tipo de criptografia abaixo.

Especificar opções de criptografia do EMRFS usando a AWS Management Console

1. Navegue até o novo console do Amazon EMR e selecione Alternar para o console antigo na navegação lateral. Para obter mais informações sobre o que esperar ao alternar para o console antigo, consulte Usar o console antigo.
2. Escolha Create cluster (Criar cluster), Go to advanced options (Ir para opções avançadas).
3. Escolha uma Release de 4.7.2 ou anterior.
4. Escolha outras opções para Software and Steps (Software e etapas), conforme apropriado para seu aplicativo e, em seguida, escolha Next (Próximo).
5. Escolha as configurações nos painéis Hardware e General Cluster Settings (Configurações gerais do cluster), conforme apropriado para seu aplicativo.
6. No painel Security (Segurança), em Authentication and encryption (Autenticação e criptografia), selecione a opção S3 Encryption (with EMRFS) (Criptografia do S3 (com EMRFS)) a ser usada.

   > **Note**
   >
   > A criptografia do lado do servidor do S3 com gerenciamento de chaves do KMS (SSE-KMS) não está disponível ao usar o Amazon EMR 4.4 ou versões anteriores.

   - Se você escolher uma opção que use Gerenciamento de chaves da AWS , escolha um ID da chave do AWS KMS. Para obter mais informações, consulte Usar AWS KMS keys para criptografia do EMRFS.
   - Se você escolher S3 client-side encryption with custom materials provider (Criptografia do lado do cliente do S3 com fornecedor de materiais personalizados), forneça o Class name (Nome da classe) e o JAR location (Local de JAR). Para obter mais informações, consulte Criptografia do lado do cliente do Amazon S3.
7. Escolha outras opções conforme apropriado para seu aplicativo e, em seguida, escolha Create Cluster (Criar cluster).

Usar AWS KMS keys para criptografia do EMRFS

A chave de AWS KMS criptografia do deve ser criada na mesma região que a instância de cluster do Amazon EMR e dos buckets do Amazon S3 usados com o EMRFS. Se a chave especificada estiver em uma conta diferente da que foi usada para configurar um cluster, será necessário especificar a chave usando o respectivo ARN.

O perfil do perfil de EC2 instância da Amazon deve ter permissões para usar a chave do KMS que você especificar. O perfil padrão para o perfil de instância no Amazon EMR é `EMR_EC2_DefaultRole`. Se você usar um perfil diferente para o perfil de instância ou usar perfis do IAM para solicitações do EMRFS para o Amazon S3, certifique-se de que cada perfil seja adicionado como um usuário de chave, conforme o caso. Isso concede ao perfil permissões para usar a chave do KMS. Para obter mais informações, consulte Using Key Policies no Guia do desenvolvedor do AWS Key Management Service e Configure IAM roles for EMRFS requests to Amazon S3.

É possível usar o AWS Management Console para adicionar seu perfil de instância ou perfil de EC2 instância à lista de usuários de chaves para a chave do KMS especificada ou usar a AWS CLI ou um AWS SDK para anexar uma política de chave apropriada.

O Amazon EMR oferece suporte somente a chaves do KMS simétricas. Não é possível usar uma chave do KMS assimétrica para criptografar dados em repouso em um cluster do Amazon EMR. Para obter ajuda para determinar se uma chave do KMS é simétrica ou assimétrica, consulte Identifying symmetric and asymmetric KMS keys.

O procedimento abaixo descreve como adicionar o perfil de instância do Amazon EMR padrão, `EMR_EC2_DefaultRole`, como um usuário de chave usando o AWS Management Console. Ele pressupõe que você já tenha criado uma chave do KMS. Para criar uma nova chave do KMS, consulte Creating Keys no Guia do desenvolvedor do AWS Key Management Service .

Para adicionar o perfil de EC2 instância do Amazon EMR à lista de usuários de chaves de criptografia

1. Faça login no console AWS Management Console e abra o AWS Key Management Service (AWS KMS) em https://console.aws.amazon.com/kms.
2. Para alterar a Região da AWS, use o seletor de regiões no canto superior direito da página.
3. Selecione o alias da chave do KMS a ser modificada.
4. Na página de detalhes da chave, em Key Users (Usuários de chaves), escolha Add (Adicionar).
5. Na caixa de diálogo Add key users (Adicionar usuários da chave) selecione a função apropriada. O nome da função padrão é `EMR_EC2_DefaultRole`.
6. Escolha Adicionar.

Criptografia do lado do servidor do Amazon S3

Quando você configura a criptografia do lado do servidor do Amazon S3, o Amazon S3 criptografa os dados no nível do objeto à medida que os grava no disco e os descriptografa quando são acessados. Para ter mais informações sobre o SSE, consulte Proteger os dados usando criptografia do lado do servidor no Guia do usuário do Amazon Simple Storage Service.

Você pode escolher entre dois sistemas de gerenciamento de chaves diferentes ao especificar a SSE no Amazon EMR:

- SSE-S3: o Amazon S3 gerencia as chaves para você.
- SSE-KMS: use uma AWS KMS key para configurar políticas apropriadas para o Amazon EMR. Para obter mais informações sobre os requisitos de chave do Amazon EMR, consulte Using AWS KMS keys for encryption.

A SSE com chaves fornecidas pelo cliente (SSE-C) não está disponível para o uso com o Amazon EMR.

Para criar um cluster com o SSE-S3 habilitado usando o AWS CLI

- Digite o seguinte comando:

```
aws emr create-cluster --release-label emr-4.7.2 or earlier \
--instance-count 3 --instance-type m5.xlarge --emrfs Encryption=ServerSide
```

Você também pode ativar o SSE-S3 configurando o fs.s3. enableServerSidePropriedade de criptografia para verdadeira nas `emrfs-site` propriedades. Consulte o exemplo para SSE-KMS a seguir e omita a propriedade para Key ID.

Para criar um cluster com o SSE-KMS ativado usando o AWS CLI

> **Note**
>
> O SSE-KMS está disponível apenas no Amazon EMR versão 4.5.0 e posteriores.

- Digite o AWS CLI comando a seguir para criar um cluster com SSE-KMS, onde *keyID* está, por exemplo AWS KMS key, um: *a4567b8-9900-12ab-1234-123a45678901*

```
aws emr create-cluster --release-label emr-4.7.2 or earlier --instance-count 3 \
--instance-type m5.xlarge --use-default-roles \
--emrfs Encryption=ServerSide,Args=[fs.s3.serverSideEncryption.kms.keyId=keyId]
```

--OU--

Digite o AWS CLI comando a seguir usando a `emrfs-site` classificação e forneça um arquivo JSON de configuração com conteúdo semelhante ao mostrado `myConfig.json` no exemplo abaixo:

```
aws emr create-cluster --release-label emr-4.7.2 or earlier --instance-count 3
 --instance-type m5.xlarge --applications Name=Hadoop --configurations file://
myConfig.json --use-default-roles
```

Exemplo de conteúdo de myConfig.json:

```
[
  {
    "Classification":"emrfs-site",
    "Properties": {
       "fs.s3.enableServerSideEncryption": "true",
       "fs.s3.serverSideEncryption.kms.keyId":"a4567b8-9900-12ab-1234-123a45678901"
    }
  }
]
```

Propriedades de configuração para SSE-S3 e SSE-KMS

Essas propriedades podem ser configuradas usando a classificação de configuração `emrfs-site`. O SSE-KMS está disponível apenas no Amazon EMR versão 4.5.0 e posteriores.

Propriedade	Valor padrão	Descrição
`fs.s3.enableServerSideEncryption`	**`false`**	Quando definida como **`true`**, objetos armazenados no Amazon S3 são criptografados usando criptografia do lado do

Propriedade	Valor padrão	Descrição
		servidor. Se nenhuma chave for especificada, SSE-S3 será usada.
`fs.s3.serverSideEncryption.kms.keyId`	**n/a**	Especifica um ID de AWS KMS chave ou ARN. Se uma chave for especificada, SSE-KMS será usada.

Criptografia do lado do cliente do Amazon S3

Com a criptografia do lado do cliente do Amazon S3, a criptografia e a descriptografia do Amazon S3 ocorrem no cliente do EMRFS em seu cluster. Os objetos são criptografados antes de serem carregados no Amazon S3 e descriptografados após serem baixados. O provedor especificado por você fornece a chave de criptografia que o cliente usa. O cliente pode usar chaves fornecidas pelo AWS KMS (CSE-KMS) ou uma classe Java personalizada que fornece a chave raiz do lado do cliente (CSE-C). As especificações de criptografia são ligeiramente diferentes entre a CSE-KMS e a CSE-C, dependendo do provedor especificado e dos metadados do objeto que está sendo descriptografado ou criptografado. Para obter mais informações sobre essas diferenças, consulte Proteger dados usando a criptografia do lado do cliente no Guia do usuário do Amazon Simple Storage Service.

> **Note**
>
> A CSE do Amazon S3 garante somente que os dados do EMRFS trocados com o Amazon S3 sejam criptografados. Não são todos os dados nos volumes de instâncias do cluster que são criptografados. Além disso, como o Hue não usa o EMRFS, os objetos que o navegador de arquivos do S3 para Hue grava no Amazon S3 não são criptografados.

Especificar CSE-KMS para dados do EMRFS no Amazon S3 usando a AWS CLI

- Digite o seguinte comando e *MyKMSKeyID* substitua pelo ID da chave ou o ARN da chave do KMS a ser usada:

```
aws emr create-cluster --release-label emr-4.7.2 or earlier
--emrfs Encryption=ClientSide,ProviderType=KMS,KMSKeyId=MyKMSKeyId
```

Criar um provedor de chaves personalizado

Dependendo do tipo de criptografia que você usa ao criar um provedor de chave personalizado, o aplicativo também deve implementar diferentes EncryptionMaterialsProvider interfaces. Ambas as as interfaces estão disponíveis na versão 1.11.0 e posteriores do AWS SDK for Java.

- Para implementar a criptografia do Amazon S3, use o com.amazonaws.sec.s3.model. EncryptionMaterialsProvider interface.
- Para implementar a criptografia de disco local, use com.amazonaws.com.amazonaws.com.amazonaws.security.elasticmapreduce.security. EncryptionMaterialsProvider interface.

Você pode usar qualquer estratégia para fornecer materiais de criptografia para a implementação. Por exemplo, você pode optar por fornecer materiais de criptografia estáticos ou fazer uma integração com um sistema de gerenciamento de chaves mais complexo.

Se estiver usando a criptografia do Amazon S3, deverá utilizar os algoritmos de criptografia AES/GCM/NoPaddingpara materiais de criptografia personalizados.

Se estiver usando criptografia de disco local, o algoritmo de criptografia a ser utilizado para materiais de criptografia personalizados varia de acordo com a versão do EMR. No Amazon EMR 7.0.0 e versões anteriores, você deve usar. AES/GCM/NoPadding No Amazon EMR 7.1.0 e versões posteriores, você deve usar AES.

A EncryptionMaterialsProvider classe obtém materiais de criptografia por contexto de criptografia. O Amazon EMR popula informações de contexto de criptografia em runtime para ajudar o chamador a determinar os materiais de criptografia corretos a serem retornados.

Example Exemplo: usar um provedor de chaves personalizado para a criptografia do Amazon S3 com o EMRFS

Quando o Amazon EMR busca materiais de criptografia da EncryptionMaterialsProvider classe para realizar a criptografia, o EMRFS opcionalmente preenche o argumento materialsDescription com dois

campos: o URI do Amazon S3 para o objeto JobFlowId e o do cluster, que podem ser usados pela classe para retornar materiais de criptografia seletivamente. EncryptionMaterialsProvider

Por exemplo, o provedor pode retornar diferentes chaves para diferentes prefixos de URI do Amazon S3. É a descrição dos materiais de criptografia retornados que acaba sendo armazenada com o objeto do Amazon S3 no lugar do valor de materialsDescription que é gerado pelo EMRFS e transmitido ao provedor. Ao descriptografar um objeto do Amazon S3, a descrição dos materiais de criptografia é transmitida à EncryptionMaterialsProvider classe, para que ela possa, mais uma vez, retornar seletivamente a chave correspondente para descriptografar esse objeto.

Uma implementação de EncryptionMaterialsProvider referência é fornecida abaixo. Outro provedor personalizado, EMRFSRSAEncryptionMaterialsProvider, está disponível em GitHub.

```
import com.amazonaws.services.s3.model.EncryptionMaterials;
import com.amazonaws.services.s3.model.EncryptionMaterialsProvider;
import com.amazonaws.services.s3.model.KMSEncryptionMaterials;
import org.apache.hadoop.conf.Configurable;
import org.apache.hadoop.conf.Configuration;

import java.util.Map;

/**
 * Provides KMSEncryptionMaterials according to Configuration
 */
public class MyEncryptionMaterialsProviders implements EncryptionMaterialsProvider,
 Configurable{
  private Configuration conf;
  private String kmsKeyId;
  private EncryptionMaterials encryptionMaterials;

  private void init() {
    this.kmsKeyId = conf.get("my.kms.key.id");
    this.encryptionMaterials = new KMSEncryptionMaterials(kmsKeyId);
  }

  @Override
  public void setConf(Configuration conf) {
    this.conf = conf;
    init();
  }

  @Override
```

```
  public Configuration getConf() {
    return this.conf;
  }

  @Override
  public void refresh() {

  }

  @Override
  public EncryptionMaterials getEncryptionMaterials(Map<String, String>
 materialsDescription) {
    return this.encryptionMaterials;
  }

  @Override
  public EncryptionMaterials getEncryptionMaterials() {
    return this.encryptionMaterials;
  }
}
```

Especificar um provedor de materiais personalizado usando a AWS CLI

Para usar a AWS CLI, transmita os argumentos `Encryption`, `ProviderType`, `CustomProviderClass` e `CustomProviderLocation` para a opção `emrfs`.

```
aws emr create-cluster --instance-type m5.xlarge --release-label emr-4.7.2 or earlier
 --emrfs Encryption=ClientSide,ProviderType=Custom,CustomProviderLocation=s3://amzn-s3-
demo-bucket/myfolder/provider.jar,CustomProviderClass=classname
```

Configurar `Encryption` como `ClientSide` permite a criptografia do lado do cliente, `CustomProviderClass` é o nome do seu objeto `EncryptionMaterialsProvider` e `CustomProviderLocation` é o local ou a localização no Amazon S3 de onde o Amazon EMR copia `CustomProviderClass` para cada nó do cluster e o coloca no caminho de classe.

Especificar um provedor de materiais personalizado usando um SDK

Para usar um SDK, primeiro você pode definir a propriedade `fs.s3.cse.encryptionMaterialsProvider.uri` para baixar a classe personalizada `EncryptionMaterialsProvider` que é armazenada no Amazon S3 em cada nó do cluster.

Isso pode ser configurado no arquivo `emrfs-site.xml`, juntamente com a CSE habilitada e a localização apropriada do provedor personalizado.

Por exemplo, no AWS SDK para Java uso RunJobFlowRequest, seu código poderá se assemelhar ao seguinte:

```
<snip>
  Map<String,String> emrfsProperties = new HashMap<String,String>();
      emrfsProperties.put("fs.s3.cse.encryptionMaterialsProvider.uri","s3://amzn-s3-
demo-bucket/MyCustomEncryptionMaterialsProvider.jar");
      emrfsProperties.put("fs.s3.cse.enabled","true");
      emrfsProperties.put("fs.s3.consistent","true");

 emrfsProperties.put("fs.s3.cse.encryptionMaterialsProvider","full.class.name.of.EncryptionMate

  Configuration myEmrfsConfig = new Configuration()
      .withClassification("emrfs-site")
      .withProperties(emrfsProperties);

  RunJobFlowRequest request = new RunJobFlowRequest()
   .withName("Custom EncryptionMaterialsProvider")
   .withReleaseLabel("emr-7.9.0")
   .withApplications(myApp)
   .withConfigurations(myEmrfsConfig)
   .withServiceRole("EMR_DefaultRole_V2")
   .withJobFlowRole("EMR_EC2_DefaultRole")
   .withLogUri("s3://myLogUri/")
   .withInstances(new JobFlowInstancesConfig()
    .withEc2KeyName("myEc2Key")
    .withInstanceCount(2)
    .withKeepJobFlowAliveWhenNoSteps(true)
    .withMasterInstanceType("m5.xlarge")
    .withSlaveInstanceType("m5.xlarge")
   );

  RunJobFlowResult result = emr.runJobFlow(request);
</snip>
```

Personalizar EncryptionMaterialsProvider com argumentos

Pode ser necessário passar argumentos diretamente para o provedor. Para fazer isso, você pode usar a classificação de configuração `emrfs-site` com argumentos personalizados definidos como

propriedades. Uma configuração de exemplo é mostrada a seguir, que é salva como um arquivo, `myConfig.json`:

```
[
    {
      "Classification": "emrfs-site",
      "Properties": {
        "myProvider.arg1":"value1",
     "myProvider.arg2":"value2"
      }
    }
 ]
```

Usando o `create-cluster` comando do AWS CLI, você pode usar a `--configurations` opção para especificar o arquivo conforme mostrado abaixo:

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge
 --instance-count 2 --configurations file://myConfig.json --emrfs
 Encryption=ClientSide,CustomProviderLocation=s3://amzn-s3-demo-bucket/myfolder/
myprovider.jar,CustomProviderClass=classname
```

Configurar o suporte ao EMRFS S3EC V2

As versões do SDK para Java no S3 (1.11.837 e posteriores) têm suporte para cliente de criptografia versão 2 (S3EC V2) com vários aprimoramentos de segurança. Para obter mais informações, consulte a publicação no blog do S3 Updates to the Amazon S3 encryption client. Além disso, consulte Amazon S3 encryption client migration no Guia do AWS SDK para Java desenvolvedor do.

O Encryption Client V1 ainda está disponível no SDK para compatibilidade retroativa. Por padrão, o EMRFS usará o S3EC V1 para criptografar e descriptografar objetos do S3 quando o CSE estiver habilitado.

Não é possível descriptografar objetos do S3 criptografados com o S3EC V2 pelo EMRFS em um cluster do EMR cuja versão seja anterior ao emr-5.31.0 (emr-5.30.1 e anteriores, emr-6.1.0 e anteriores).

Example Configurar o EMRFS para usar o S3EC V2

Para configurar o EMRFS para usar o S3EC V2, adicione esta configuração:

```
{
```

```
  "Classification": "emrfs-site",
  "Properties": {
    "fs.s3.cse.encryptionV2.enabled": "true"
  }
}
```

Propriedades **`emrfs-site.xml`** para criptografia do lado do cliente do Amazon S3

Propriedade	Valor padrão	Descrição
`fs.s3.cse.enabled`	**`false`**	Quando definido como **`true`**, os objetos do EMRFS armazenados no Amazon S3 são criptografados usando a criptografia do lado do cliente.
`fs.s3.cse.encryptionV2.enabled`	**`false`**	Quando estiver definido como `true`, o EMRFS usará o cliente de criptografia do S3 versão 2 para criptografar e descriptografar objetos no S3. Disponível no EMR 5.31.0 e versões posteriores.
`fs.s3.cse.encryptionMaterialsProvider.uri`	**`N/A`**	Aplica-se ao usar materiais de criptografia personalizada. O URI do Amazon S3 em que o JAR com `EncryptionMaterialsProvider` está localizado. Quando você fornecer esse URI, o Amazon EMR baixará o JAR automaticamente em todos os nós do cluster.

Propriedade	Valor padrão	Descrição
`fs.s3.cse.encryptionMaterialsProvider`	**N/A**	O caminho da classe `EncryptionMaterialsProvider` usado com a criptografia no lado do cliente. Ao usar CSE-KMS, especifique `com.amazon.ws.emr.hadoop.fs.cse.KMSEncryptionMaterialsProvider`.
`fs.s3.cse.materialsDescription.enabled`	**false**	Quando definido como `true`, preenche a MaterialsDescription dos objetos criptografados com o URI do Amazon S3 para o objeto e o. JobFlowId Defina como `true` ao usar materiais de criptografia personalizados.
`fs.s3.cse.kms.keyId`	**N/A**	Aplica-se ao usar a CSE-KMS. O valor do KeyId ARN ou alias da chave do KMS usados para criptografia.

Propriedade	Valor padrão	Descrição
`fs.s3.cse.cryptoStorageMode`	**`ObjectMetadata`**	O modo de armazenamento do Amazon S3. Por padrão, a descrição das informações de criptografia é armazenada nos metadados do objeto. Você também pode armazenar a descrição em um arquivo de instruções. Os valores válidos são ObjectMetadata InstructionFile e. Para obter mais informações, consulte Client-side data encryption with the AWS SDK para Java and Amazon S3.

Criptografia do lado do cliente Amazon S3 com S3A

A partir da versão 7.6.0 do Amazon Elastic Map Reduce (EMR), o conector do sistema de arquivos S3A agora oferece suporte à criptografia do lado do cliente do Amazon S3. Isso significa que a criptografia e a descriptografia dos dados do Amazon S3 ocorrem diretamente no cliente S3A em seu cluster de computação. Ao usar esse recurso, os arquivos são criptografados automaticamente antes de serem enviados para o Amazon S3 e descriptografados no momento do download. Para obter detalhes abrangentes sobre a metodologia de criptografia e sua implementação, os usuários podem consultar Proteção de dados usando criptografia do lado do cliente no Guia do usuário do Amazon Simple Storage Service.

Ao habilitar a criptografia do lado do cliente (CSE) com o S3A no Amazon EMR, você tem duas opções de sistema de gerenciamento de chaves:

- CSE-KMS — Essa abordagem utiliza uma chave AWS Key Management Service (KMS) configurada com políticas projetadas especificamente para o Amazon EMR. Para obter informações detalhadas sobre os principais requisitos, consulte a documentação Usando AWS KMS chaves para criptografia.
- CSE-CUSTOM — Esse método permite integrar uma classe Java personalizada que fornece a chave raiz do lado do cliente responsável por criptografar e descriptografar dados.

> **Note**
>
> A criptografia do lado do cliente S3A no EMR é inerentemente compatível com a criptografia do lado do cliente do EMRFS, o que significa que objetos criptografados usando o EMRFS CSE podem ser lidos por meio do S3A CSE.

Tópicos

- Configurar CSE-KMS
- Configurar CSE-CUSTOM
- Propriedades da criptografia do lado do cliente do Amazon S3 com S3A

Configurar CSE-KMS

Você pode habilitar a criptografia do lado do cliente usando AWS KMS (CSE-KMS) em dois escopos principais:

- A primeira é a configuração em todo o cluster:

```
[
  {
    "Classification":"core-site",
    "Properties": {
       "fs.s3a.encryption.algorithm": "CSE-KMS",
       "fs.s3a.encryption.key":"${KMS_KEY_ID}",
    }
  }
]
```

> **Note**
>
> Se a região da AWS KMS chave for diferente da região S3 Bucket/EMR, você deverá definir a seguinte configuração adicional:. `fs.s3a.encryption.cse.kms.region=${KMS_REGION}`

- A segunda é a configuração específica do trabalho ou do aplicativo. O CSE-KMS pode ser configurado para um aplicativo Spark específico da seguinte forma:

```
spark-submit --conf spark.hadoop.fs.s3a.encryption.algorithm=CSE-KMS --conf
 spark.hadoop.fs.s3a.encryption.key=${KMS_KEY_ID}
```

Configurar CSE-CUSTOM

Para usar o CSE-CUSTOM, você deve criar um provedor de chave personalizado implementando a interface do Keyring. Aqui está um exemplo de implementação:

```
public class CustomKeyring implements Keyring {
  public CustomKeyring()  {
    // custom code
  }
```

```
  @Override
  public EncryptionMaterials onEncrypt(EncryptionMaterials encryptionMaterials) {
    // custom code
  }

  @Override
  public DecryptionMaterials onDecrypt(DecryptionMaterials decryptionMaterials,
      ListEncryptedDataKey list) {
    // custom code
  }
```

Você pode ativar as chaves personalizadas de criptografia do lado do cliente (CSE-CUSTOM) em dois escopos principais:

- O primeiro escopo é a configuração em todo o cluster:

```
[
  {
    "Classification":"core-site",
    "Properties": {
       "fs.s3a.encryption.algorithm": "CSE-CUSTOM",
       "fs.s3a.cse.customKeyringProvider.uri":"S3 path of custom jar",
       "fs.s3a.encryption.cse.custom.keyring.class.name":"fully qualified class name"
    }
  }
]
```

- A segunda é a configuração específica do trabalho ou do aplicativo. O CSE-CUSTOM pode ser configurado para um aplicativo Spark específico da seguinte forma:

```
spark-submit --conf spark.hadoop.fs.s3a.encryption.algorithm=CSE-CUSTOM --conf
 spark.hadoop.fs.s3a.encryption.cse.custom.keyring.class.name=fully qualified class
 name
```

> **Note**
>
> Certifique-se de que o jar personalizado necessário para gerar chaves de criptografia/descriptografia esteja presente no caminho da classe.

Propriedades da criptografia do lado do cliente do Amazon S3 com S3A

Para configurar a criptografia do lado do cliente com o S3A, há várias propriedades de configuração que devem ser definidas nas configurações do core-site.xml. Para obter mais informações sobre configurações personalizadas, consulte Configurar aplicativos.

Propriedade	Valor padrão	Descrição
fs.s3a.encryption.algorithm	N/A	Quando definido como CSE-KMS ou CSE-CUSTOM , os objetos armazenados no Amazon S3 são criptografados usando criptografia do lado do cliente.
fs.s3a.encryption.key	N/A	Aplica-se ao usar CSE-KMS. O valor do KeyId, ARN ou alias da chave KMS usada para criptografia.
fs.s3a.encryption.cse.kms.region	N/A	Aplica-se ao usar CSE-KMS. A região em que a AWS KMS chave é gerada. Por padrão, a região KMS é definida com valores iguais aos da região de cluster S3 Bucket/EMR.
fs.s3a.encryption.cse.custom.keyring.class.name	N/A	Aplica-se ao usar CSE-KMS. O nome da classe totalmente qualificado do provedor de chave personalizada.
fs.s3a.cse. customKeyringProvider.uri	N/A	Aplica-se ao usar CSE-CUSTOM . O URI do Amazon S3 em que o JAR com a implementação personalizada

Propriedade	Valor padrão	Descrição
		do Keyring está localizado. Quando você fornecer esse URI, o Amazon EMR baixará o JAR automaticamente em todos os nós do cluster.
fs.s3a.encryption.cse.v1.compatibility.enabled	`'true'`	Isso fornece compatibilidade com versões anteriores de clientes SDK mais antigos, como o usado com o EMRFS. Desative isso, quando não houver essa dependência, para obter um melhor desempenho.

CloudWatch Agente da Amazon

O CloudWatch agente da Amazon no Amazon EMR é uma ferramenta capaz de monitorar as EC2 instâncias da Amazon no cluster do EMR. É possível armazenar e visualizar as métricas que você coletar com o CloudWatch agente em CloudWatch. Para obter mais informações sobre o CloudWatch agente, consulte o Guia CloudWatch do usuário da Amazon.

> **Note**
>
> Você receberá cobranças adicionais se usar outros Serviços da AWS para publicar, consultar ou visualizar as métricas dos CloudWatch agentes da Amazon. Consulte as páginas a seguir para obter mais informações sobre preços.
>
> - CloudWatch Preços da Amazon
> - Preços do Amazon Managed Service para Prometheus
> - Preços do Amazon Managed Grafana
>
> Essas métricas são separadas das métricas gratuitas que o Amazon EMR fornece sob o `AWS/ElasticMapReduce` namespace em. CloudWatch Para obter mais informações sobre as métricas que não são publicadas pelo CloudWatch agente, consulte Monitoring Amazon EMR metrics with. CloudWatch

Com o Amazon EMR 7.0 e superior, você pode instalar uma versão personalizada do CloudWatch agente da Amazon EMR no cluster para coletar métricas do cluster do EMR. CloudWatchAgent é compatível com Runtime Role Clusters para EMR 7.6 e versões posteriores. O agente coleta as métricas de todos os nós do cluster, reúne-as no nó primário e publica as métricas desse nó na nuvem.

Você pode instalar o agente ao criar um cluster no console ou ao usar a API `create-cluster`. Para obter mais informações, consulte Criação de um cluster do EMR que use o agente da Amazon CloudWatch . Por padrão, um cluster do Amazon EMR que executa a série da versão 7.x publica 34 métricas CloudWatch em nível de sistema em intervalos de 60 segundos, mas você pode configurar o agente para publicar métricas diferentes. Outra opção é publicar métricas no Amazon Managed Service para Prometheus, e você também pode escolher quais métricas publicar. Em diferentes casos de uso e configurações, você pode configurar o agente para exibir e consultar as métricas

no CloudWatch console, no Amazon Managed Grafana ou por meio do for CloudWatch ou APIs do Amazon Managed Service for Prometheus. Esses Serviços da AWS geram cobranças quando você os usa para armazenar e consultar métricas.

O CloudWatch agente no Amazon EMR pode publicar métricas do sistema, bem como métricas do JMX para os serviços a seguir no cluster do Amazon EMR.

- Hadoop DataNode
- Hadoop NameNode
- Fio NodeManager
- Fio ResourceManager
- HBase Master: somente Amazon EMR 7.1+
- HBase RegionServer : somente Amazon EMR 7.1+
- HBase ThriftServer : somente Amazon EMR 7.1+

Para obter mais informações sobre métricas disponíveis e como configurar o CloudWatch agente no Amazon EMR, consulte Configuring agent agent CloudWatch for Amazon EMR.

A tabela a seguir lista a versão AmazonCloudWatchAgent incluída na versão mais recente da série 7.x do Amazon EMR, além dos componentes com os quais o Amazon EMR é instalado. AmazonCloudWatchAgent

Para obter a versão dos componentes instalados AmazonCloudWatchAgent nessa versão, consulte Release 7.9.0 Component Versions.

AmazonCloudWatchAgent informações da versão 7.9.0

Rótulo de versão do Amazon EMR	AmazonCloudWatchAgent Versão	Componentes instalados com o AmazonCloudWatchAgent
emzn-7.9.0	AmazonCloudWatchAgent 1.300032.2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent , emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-nameno

Rótulo de versão do Amazon EMR	AmazonCloudWatchAgent Versão	Componentes instalados com o AmazonCloudWatchAgent
		de, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Tópicos

- Criação de um cluster do EMR que use o agente da Amazon CloudWatch
- Métricas padrão do CloudWatch agente com o Amazon EMR
- Configuração do CloudWatch agente para o Amazon EMR
- Considerações e limitações
- CloudWatch histórico de versões do agente

Criação de um cluster do EMR que use o agente da Amazon CloudWatch

Os procedimentos nesta seção descrevem as etapas para criar um cluster no Amazon EMR com o CloudWatch agente da Amazon do AWS Management Console e da. AWS CLI

Tópicos

- Permissões necessárias do IAM para o CloudWatch agente
- Endpoint obrigatório do CloudWatch agente
- Criação de um cluster do EMR

Permissões necessárias do IAM para o CloudWatch agente

O CloudWatch agente exige a `cloudwatch:PutMetricData` permissão AWS Identity and Access Management (IAM) no perfil de EC2 instância da Amazon para o Amazon EMR. O perfil padrão do Amazon EMR já tem essa permissão. Você pode criar a função padrão a partir do AWS CLI with`aws`

`emr create-default-roles`. Para obter mais informações, consulte Função de serviço para EC2 instâncias de cluster (perfil de EC2 instância) no Guia de gerenciamento do Amazon EMR.

O seguinte exemplo de política do IAM inclui a permissão `cloudwatch:PutMetricData`:

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "cloudwatch:PutMetricData",
      "Resource": "*"
    }
  ]
}
```

Endpoint obrigatório do CloudWatch agente

Para publicar métricas CloudWatch para um cluster do EMR em uma sub-rede privada, crie um endpoint do CloudWatch agente e associe-o à VPC na qual a sub-rede privada está.

Para obter mais informações sobre os CloudWatch endpoints de cada um Região da AWS, consulte os CloudWatch endpoints e cotas da Amazon no Guia de referência AWS geral.

Criação de um cluster do EMR

Depois de configurar as permissões e o endpoint necessários para usar com o CloudWatch agente, use o AWS Management Console ou a AWS CLI para criar um cluster com o agente instalado.

Console

Para criar um cluster com o CloudWatch agente da Amazon no console

1. Navegue até o console do Amazon EMR.
2. Selecione Criar cluster.
3. Em Nome e aplicações, escolha uma versão 7.0.0 ou superior do Amazon EMR.
4. Em Pacote de aplicações, selecione o pacote ou as aplicações que deseja instalar no cluster e inclua CloudWatch Agente nas seleções.
5. Prossiga com a criação do cluster para atender às necessidades do seu caso de uso.

AWS CLI

No AWS CLI, você pode adicionar o CloudWatch agente da Amazon a um cluster com o `--applications` parâmetro para`create-cluster`.

Para criar um cluster com o CloudWatch agente da Amazon a partir do AWS CLI

- Ao criar um cluster, use um comando semelhante ao mostrado a seguir para incluir o CloudWatch agente da Amazon. *myKey*Substitua pelo nome do par de EC2 chaves.

```
aws emr create-cluster --name "Spark cluster with CloudWatch agent" \
--release-label emr-7.0.0 \
--applications Name=Spark Name=AmazonCloudWatchAgent \
--ec2-attributes KeyName=myKey --instance-type m7g.2xlarge \
--instance-count 3 --use-default-roles
```

Para obter mais detalhes sobre como usar o Amazon EMR com a AWS CLI, consulte a Referência de AWS CLI comandos da.

Métricas padrão do CloudWatch agente com o Amazon EMR

Ao instalar o CloudWatch agente da Amazon EMR, a configuração padrão publica as métricas a seguir do sistema para todas as instâncias no cluster, a menos que você configure o agente de forma diferente. Para obter as definições de cada métrica, consulte Métricas coletadas pelo CloudWatch agente no Guia CloudWatch do usuário da Amazon.

CPU

Métricas de CPU

- `cpu_usage_active`
- `cpu_usage_guest`
- `cpu_usage_guest_nice`
- `cpu_usage_idle`
- `cpu_usage_iowait`
- `cpu_usage_irq`
- `cpu_usage_nice`

- cpu_usage_softirq
- cpu_usage_steal
- cpu_usage_system
- cpu_usage_user

Disk

Métricas de disco

- disk_free
- disk_total
- disk_used
- disk_used_percent

Memory

Métricas de memória

- mem_active
- mem_available
- mem_available_percent
- mem_free
- mem_inactive
- mem_total
- mem_used
- mem_used_percent
- mem_buffered
- mem_cached

Network IO

Métricas de E/S de rede

- net_bytes_recv

- net_bytes_sent
- net_packets_recv
- net_packets_sent

Process

Métricas de processo

- processes_running
- processes_total

Swap

Métricas de troca

- swap_free
- swap_used
- swap_used_percent

Por padrão, o agente publica todas as CloudWatch métricas no namespace personalizado e no esquema CWAgentcluster.id,,,. instance.id node.type service.name Use as seguintes etapas para visualizar essas métricas padrão que o CloudWatch agente publica para o Amazon EMR:

1. Navegue até o CloudWatch console da Amazon.
2. Escolha a guia Métricas e Todas as métricas.
3. Em Namespaces personalizados, selecione **CWAgent**. Em seguida, selecione o esquema: **cluster.id**, **instance.id**, **node.type**, **service.name**.
4. Continue consultando as métricas da CloudWatch interface conforme exigido pelo seu caso de uso.

O Amazon EMR anexa os seguintes rótulos a cada métrica: ["cluster.id", "instance.id', "node.type", "service.name"]. Considere o seguinte em relação a esses rótulos:

- O valor do rótulo cluster.id é o ID do cluster do EMR que produziu a métrica. Um exemplo de valor para o rótulo cluster.id é j-123456789ABC.

- O valor do rótulo `instance.id` é o ID da instância no cluster do EMR que produziu a métrica. Um exemplo de valor para o rótulo `instance.id` é `i-01bcf5f140f3355777`.
- O valor do rótulo `node.type` representa o tipo de nó da instância no cluster que produziu a métrica. O valor pode ser `master`, `core` ou `task`.
- O valor do rótulo `service.name` é o nome do serviço que produziu a métrica. O valor `service.name` padrão para as métricas padrão do sistema é `system`.

> **Note**
>
> A versão 7.0.0 do Amazon EMR usa um esquema de rótulos de métricas um pouco diferente. `cluster.id` é `jobflow.id` e `node.type` não está disponível. `instance.id` e `service.name` estão disponíveis.

Configuração do CloudWatch agente para o Amazon EMR

O Amazon EMR 7.0.0 e versões posteriores incluem o agente da Amazon EMR 7.0.0. CloudWatch As seções a seguir descrevem como configurar o agente no cluster do EMR, conforme descrito em Criação de um cluster do EMR que use o agente da Amazon CloudWatch .

Tópicos

- Configuração do CloudWatch agente para o Amazon EMR 7.1.0
- Configuração do CloudWatch agente para o Amazon EMR 7.0.0

Configuração do CloudWatch agente para o Amazon EMR 7.1.0

A partir do Amazon EMR 7.1.0, você pode configurar o CloudWatch agente da Amazon EMR para usar métricas adicionais do sistema, adicionar métricas de aplicações e alterar o destino das métricas usando a API de configuração do Amazon EMR. Para obter mais informações sobre como usar a API de configuração do EMR para configurar as aplicações do cluster, consulte Configure applications.

> **Note**
>
> A versão 7.1.0 é compatível apenas com o tipo de reconfiguração `OVERWRITE`. Para obter mais informações sobre os tipos de reconfiguração, consulte Considerations when you reconfigure an instance group.

Tópicos

- Esquema de configuração
- Exemplos de configurações de métricas do sistema
- Exemplos de configurações de métricas de aplicações
- Exemplo do Amazon Managed Service para Prometheus

Esquema de configuração

`emr-metrics` tem as seguintes classificações:

- `emr-system-metrics`: configure métricas do sistema, como CPU, disco e memória.
- `emr-hadoop-hdfs-datanode-metrics`— configurar métricas do Hadoop DataNode JMX
- `emr-hadoop-hdfs-namenode-metrics`— configurar métricas do Hadoop NameNode JMX
- `emr-hadoop-yarn-nodemanager-metrics`— configurar métricas do Yarn NodeManager JMX
- `emr-hadoop-yarn-resourcemanager-metrics`— configurar métricas do Yarn ResourceManager JMX
- `emr-hbase-master-metrics`— configurar métricas do HBase Master JMX
- `emr-hbase-region-server-metrics`— configurar métricas JMX do HBase Region Server
- `emr-hbase-rest-server-metrics`— configurar métricas JMX do servidor HBase REST
- `emr-hbase-thrift-server-metrics`— configurar métricas JMX do HBase Thrift Server

As tabelas a seguir descrevem as propriedades e configurações disponíveis para todas as classificações.

propriedades do emr-metrics

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	Observações
`metrics_destination`	Opcional	Determina se as métricas do cluster são publicadas na Amazon CloudWatch ou no Amazon Managed Service for Prometheus.	"CLOUDWATCH"	"CLOUDWATCH", "PROMETHEUS"	Essa propriedade não diferencia maiúsculas de minúsculas. Por exemplo, "Cloudwatch" é o mesmo que "CLOUDWATCH".
`prometheus_endpoint`	Opcional	Se `metrics_destination` estiver definido como “PROMETHEUS”, essa propriedade configura o CloudWatch agente para enviar métricas ao endpoint de gravação remota do Amazon Managed Service para	N/D	Qualquer URL de gravação remota do Amazon Managed Service para Prometheus. O formato do URL de gravação remota é `https://aps-workspaces.<region>.com/workspaces/<workspace_id>/`	Esse campo será obrigatório se `metrics_destination` estiver definido como "PROMETHEUS". O provisionamento falhará se você não fornecer uma chave ou se o valor for uma string vazia.

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	Observações
		Prometheus fornecido.		`api/v1/remote_write`	

emr-system-metrics propriedades

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	Observações
`metrics_collection_interval`	Opcional	Com que frequência, em segundos, as métricas são coletadas e publicadas pelo CloudWatch agente.	"60"	Uma string especificando o número de segundos. Só aceita números inteiros.	Você pode substituir essa propriedade pela propriedade `metrics_collection_interval` de grupos de métricas individuais.

emr-system-metrics configurações

cpu

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	Observações
`metrics`	Opcional	A lista de métricas de CPU para	Consulte Métricas padrão do CloudWatch	Uma lista separada por vírgula de nomes de	Especificar uma string vazia significa

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	Observações
		o agente coletar.	agente com o Amazon EMR	métricas de CPU válidos com ou sem o prefixo `cpu_`, como `usage_active` e `cpu_time_idle`. Consulte Métricas coletadas pelo CloudWatch agente para obter métricas válidas.	não publicar nenhuma métrica de CPU.
`metrics_collection_interval`	Opcional	Com que frequência, em segundos, o agente deve coletar e publicar métricas de CPU.	O valor do `metrics_collection_interval` global.	Uma string especificando o número de segundos. Aceita somente números inteiros.	Esse valor substitui a propriedade global `metrics_collection_interval` somente para métricas de CPU.

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	Observações
`drop_original_metrics`	Opcional	Lista de métricas de CPU para as quais não publicar métricas não agregadas.	Nenhuma métrica de CPU não agregada publicada.	Uma lista separada por vírgulas de métricas de CPU que também são especificadas na propriedade de métricas. Uma string vazia significa publicar todas as métricas de CPU.	O CloudWatch agente agrega todas as métricas por ID do cluster, ID da instância, tipo de nó e nome do serviço. Por padrão, o CloudWatch agente não publica as métricas por recurso para métricas com vários recursos.
`resources`	Opcional	Determina se o agente publicará métricas por núcleo.	"*"	"*" habilita métricas por núcleo. "" desabilita métricas por núcleo.	O CloudWatch agente publica somente métricas por núcleo para métricas de CPU que não são descartadas. `drop_original_metrics`

disk

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	Observações
`metrics`	Opcional	A lista de métricas de disco para o agente coletar.	Consulte Métricas padrão do CloudWatch agente com o Amazon EMR	Uma lista separada por vírgula de nomes de métricas de disco válidos com ou sem o prefixo `disk_`, como `disk_total` e `used_percent`. Consulte Métricas coletadas pelo CloudWatch agente para obter métricas válidas.	Especificar uma string vazia significa não publicar nenhuma métrica de disco.
`metrics_collection_interval`	Opcional	Com que frequência, em segundos, o agente deve coletar e publicar	O valor do `metrics_collection_interval` global.	Uma string especificando o número de segundos. Aceita somente	Esse valor substitui a propriedade global `metrics_collection_interval` somente

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	Observações
		métricas de disco.		números inteiros.	para métricas de disco.
`drop_original_metrics`	Opcional	Lista de métricas de disco para as quais não publicar métricas não agregadas.	Nenhuma métrica de disco não agregada publicada.	Uma lista separada por vírgulas de métricas de disco que também são especificadas na propriedade de métricas. Uma string vazia significa publicar todas as métricas do disco.	O CloudWatch agente agrega todas as métricas por ID do cluster, ID da instância, tipo de nó e nome do serviço. Por padrão, o CloudWatch agente não publica as métricas por recurso para métricas com vários recursos.

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	Observações
`resources`	Opcional	Determina se o agente publicará per-mount-point métricas.	"*"	"*" significa todos os pontos de montagem, "" significa nenhum ponto de montagem ou uma lista separada por vírgulas de pontos de montagem. Por exemplo, ."/,,emr"	O CloudWatch agente publica somente per-mount-point métricas para métricas de disco que não são descartadas `drop_original_metrics` .

diskio

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	Observações
`metrics`	Opcional	A lista de métricas de E/S do disco para o agente coletar.	Consulte Métricas padrão do CloudWatch agente com o Amazon EMR	Uma lista separada por vírgulas de nomes de métricas de E/S de disco válidos com ou sem o prefixo `diskio_`, como	Especificar uma string vazia significa não publicar nenhuma métrica de E/S de disco.

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	Observações
				`diskio_reads` e `writes`. Consulte Métricas coletadas pelo CloudWatch agente para obter métricas válidas.	
`metrics_collection_interval`	Opcional	Com que frequência, em segundos, o agente deve coletar e publicar métricas de E/S do disco.	O valor do `metrics_collection_interval` global.	Uma string especificando o número de segundos. Aceita somente números inteiros.	Esse valor substitui a propriedade global `metrics_collection_interval` somente para métricas de E/S do disco.

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	Observações
`drop_original_metrics`	Opcional	Lista de métricas de E/S de disco para as quais não publicar métricas não agregadas.	Nenhuma métrica de E/S de disco não agregada publicada.	Uma lista separada por vírgulas de métricas de E/S de disco que também são especificadas na propriedade de métricas. Uma string vazia significa publicar todas as métricas de E/S do disco.	O CloudWatch agente agrega todas as métricas por ID do cluster, ID da instância, tipo de nó e nome do serviço. Por padrão, o CloudWatch agente não publica as métricas por recurso para métricas com vários recursos.

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	Observações
`resources`	Opcional	Determina se o agente publicará métricas por dispositivo.	"*"	"*" significa todos os dispositivos de armazenamento, "" significa nenhum dispositivo de armazenamento ou uma lista de nomes de dispositivos separados por vírgulas. Por exemplo, ."nvr nvme1n1"	O CloudWatch agente publica somente métricas por dispositivo para métricas de E/S de disco que não são descartadas. `drop_original_metrics`

mem

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	Observações
`metrics`	Opcional	A lista de métricas de memória para o agente coletar.	Consulte Métricas padrão do CloudWatch agente com o Amazon EMR	Uma lista separada por vírgula de nomes de métricas de memória válidos com	Especificar uma string vazia significa não publicar nenhuma

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	Observações
				ou sem o prefixo `mem_`, como `mem_available` e `available_percent`. Consulte Métricas coletadas pelo CloudWatch agente para obter métricas válidas.	métrica de memória.
`metrics_collection_interval`	Opcional	Com que frequência, em segundos, o agente deve coletar e publicar métricas de memória.	O valor do `metrics_collection_interval` global.	Uma string especificando o número de segundos. Aceita somente números inteiros.	Esse valor substitui a propriedade global `metrics_collection_interval` somente para métricas de memória.

net

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	Observações
`metrics`	Opcional	A lista de métricas de rede para o agente coletar.	Consulte Métricas padrão do CloudWatch agente com o Amazon EMR	Uma lista separada por vírgula de nomes de métricas de rede válidos com ou sem o prefixo `net_`, como `net_packets_sent` e `packets_recv`. Consulte Métricas coletadas pelo CloudWatch agente para obter métricas válidas.	Especificar uma string vazia significa não publicar nenhuma métrica de rede.
`metrics_collection_interval`	Opcional	Com que frequência, em segundos, o agente deve coletar e publicar métricas de rede.	O valor do `metrics_collection_interval` global.	Uma string especificando o número de segundos. Aceita somente números inteiros.	Esse valor substitui a propriedade global `metrics_collection_interval` somente para

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	Observações
					métricas de rede.
`drop_original_metrics`	Opcional	Lista de métricas de rede para as quais não publicar métricas não agregadas.	Nenhuma métrica de rede não agregada publicada.	Uma lista separada por vírgulas de métricas de rede que também são especificadas na propriedade de métricas. Uma string vazia significa publicar todas as métricas da rede.	O CloudWatch agente agrega todas as métricas por ID do cluster, ID da instância, tipo de nó e nome do serviço. Por padrão, o CloudWatch agente não publica as métricas por recurso para métricas com vários recursos.

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	Observações
`resources`	Opcional	Determina se o agente publicará métricas por interface.	"*"	"*" significa todas as interfaces de rede, "" significa nenhuma interface de rede ou uma lista separada por vírgulas de nomes de interfaces. Por exemplo, `."eth1"`	O CloudWatch agente publica somente métricas por interface para métricas de rede que não são descartadas. `drop_original_metrics`

netstat

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	Observações
`metrics`	Opcional	A lista de métricas de estatísticas de rede para o agente coletar.	Consulte Métricas padrão do CloudWatch agente com o Amazon EMR	Uma lista separada por vírgula de nomes de métricas de memória válidos com ou sem o prefixo `netstat_`, como	Especificar uma string vazia significa não publicar nenhuma métrica de estatística de rede.

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	Observações
				`tcp_listen` e `netstat_udp_socket`. Consulte Métricas coletadas pelo CloudWatch agente para obter métricas válidas.	
`metrics_collection_interval`	Opcional	Com que frequência, em segundos, o agente deve coletar e publicar métricas de estatística de rede.	O valor do `metrics_collection_interval` global.	Uma string especificando o número de segundos. Aceita somente números inteiros.	Esse valor substitui a propriedade global `metrics_collection_interval` somente para métricas de estatística de rede.

processes

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	Observações
`metrics`	Opcional	A lista de métricas do processo que o agente deve coletar.	Consulte Métricas padrão do CloudWatch agente com o Amazon EMR	Uma lista separada por vírgula de nomes de métricas de memória válidos com ou sem o prefixo `processes_` , como `processes_running` e `total`. Consulte Métricas coletadas pelo CloudWatch agente para obter métricas válidas.	Especificar uma string vazia significa não publicar nenhuma métrica do processo.
`metrics_collection_interval`	Opcional	Com que frequência, em segundos, o agente deve coletar e publicar as métricas do	O valor do `metrics_collection_interval` global.	Uma string especificando o número de segundos. Aceita somente	Esse valor substitui a propriedade global `metrics_collection_interval` somente

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	Observações
		processo do sistema.		números inteiros.	para métricas de processo do sistema.

swap

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	Observações
`metrics`	Opcional	A lista de métricas de troca que o agente deve coletar.	Consulte Métricas padrão do CloudWatch agente com o Amazon EMR	Uma lista separada por vírgula de nomes de métricas de memória válidos com ou sem o prefixo `swap_`, como `swap_free` e `used_percent` . Consulte Métricas coletadas pelo CloudWatch agente para obter	Especificar uma string vazia significa não publicar nenhuma métrica de troca.

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	Observações
				métricas válidas.	
`metrics_collection_interval`	Opcional	Com que frequência, em segundos, o agente deve coletar e publicar métricas de troca.	O valor do `metrics_collection_interval` global.	Uma string especificando o número de segundos. Aceita somente números inteiros.	Esse valor substitui a propriedade global `metrics_collection_interval` somente para métricas de troca.

emr-hadoop-hdfs-datanode-propriedades métricas

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	
<custom_bean_name>	Opcional	N/D	O CloudWatch agente MBean desse agente deve coletar métricas, como`Hadoop:service=DataNode,name=DataNodeActivity` . Você encontra	Uma string contendo a lista delimitada por vírgulas de métricas associadas ao. MBean Por exemplo, `.BlocksCached,BlocksRead`	

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	
			exemplos de MBean nomes e suas métricas correspondentes nos exemplos de arquivos YAML do JMX para o Amazon EMR versão 7.0.		
`otel.metric.export.interval`	Opcional	Com que frequência, em milissegundos, coletar métricas do Hadoop DataNode .	"60000"	Uma string especificando o número de milissegundos. Aceita somente números inteiros.	

emr-hadoop-hdfs-namenode-propriedades métricas

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	
<custom_bean_name>	Opcional	N/D	O CloudWatch agente MBean desse agente deve coletar métricas, comoHadoop:s	Uma string contendo a lista delimitada por vírgulas de métricas associadas ao.	

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	
			`rvice=NameNode,name=FSNamesystem` . Você encontra exemplos de MBean nomes e suas métricas correspondentes nos exemplos de arquivos YAML do JMX para o Amazon EMR versão 7.0.	MBean Por exemplo, `.BlockCapacity,CapacityUsedGB`	
`otel.metric.export.interval`	Opcional	Com que frequência, em milissegundos, coletar métricas do Hadoop NameNode .	"60000"	Uma string especificando o número de milissegundos. Aceita somente números inteiros.	

emr-hadoop-yarn-nodemanager-propriedades métricas

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	
<custom_bean_name>	Opcional	N/D	O CloudWatch agente MBean desse agente deve coletar métricas, como`Hadoop:service=NodeManager,name=NodeManagerMetrics` . Você encontra exemplos de MBean nomes e suas métricas correspondentes nos exemplos de arquivos YAML do JMX para o Amazon EMR versão 7.0.	Uma string contendo a lista delimitada por vírgulas de métricas associadas ao. MBean Por exemplo, `.MaxCapacity,AllocatedGB`	
`otel.metric.export.interval`	Opcional	Com que frequência, em milissegundos, coletar métricas YARN do	"60000"	Uma string especificando o número de milissegundos. Aceita somente	

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	
		Hadoop. NodeManager		números inteiros.	

emr-hadoop-yarn-resourcemanager-propriedades métricas

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	
<custom_bean_name>	Opcional	N/D	O CloudWatch agente MBean desse agente deve coletar métricas, comoHadoop:service=ResourceManager,name=PartitionQueueMetrics . Você encontra exemplos de MBean nomes e suas métricas correspondentes nos exemplos de arquivos YAML do JMX para o	Uma string contendo a lista delimitada por vírgulas de métricas associadas ao. MBean Por exemplo, .MaxCapacity,MaxCapacityVCores	

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	
			Amazon EMR versão 7.0.		
`otel.metric.export.interval`	Opcional	Com que frequência, em milissegundos, coletar métricas YARN do Hadoop. ResourceManager	"60000"	Uma string especificando o número de milissegundos. Aceita somente números inteiros.	

emr-hbase-master-metrics propriedades

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	
<custom_bean_name>	Opcional	N/D	O CloudWatch agente MBean desse agente deve coletar métricas, como`Hadoop:service=HBase,name=Master,sub=AssignmentManager` . Você encontra exemplos	Uma string contendo a lista delimitada por vírgulas de métricas associadas ao. MBean Por exemplo, `.AssignFailedCount, AssignSubmittedCount`	

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	
			de MBean nomes e suas métricas correspon dentes nos exemplos de arquivos YAML do JMX para o Amazon EMR versão 7.0.		
`otel.metric.export.interval`	Opcional	Com que frequência, em milissegundos, coletar métricas do HBase mestre.	"60000"	Uma string especificando o número de milissegundos. Aceita somente números inteiros.	

emr-hbase-region-server-propriedades métricas

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	
<custom_bean_name>	Opcional	N/D	O CloudWatch agente MBean desse agente deve coletar métricas, como`Hadoop:service=HBa`	Uma string contendo a lista delimitada por vírgulas de métricas associadas ao. MBean Por

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	
			`se,name=RegionServer,sub=IPC` . Você encontra exemplos de MBean nomes e suas métricas correspondentes nos exemplos de arquivos YAML do JMX para o Amazon EMR versão 7.0.	exemplo, `.numActiveHandler,numActivePriorityHandler`	
`otel.metric.export.interval`	Opcional	Com que frequência, em milissegundos, coletar métricas do servidor de HBase regiões.	"60000"	Uma string especificando o número de milissegundos. Aceita somente números inteiros.	

emr-hbase-rest-server-propriedades métricas

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores
<custom_bean_name>	Opcional	N/D	O CloudWatch agente MBean desse agente deve coletar métricas, como`Hadoop:service=HBase,name=REST` . Você encontra exemplos de MBean nomes e suas métricas correspondentes nos exemplos de arquivos YAML do JMX para o Amazon EMR versão 7.0.	Uma string contendo a lista delimitada por vírgulas de métricas associadas ao. MBean Por exemplo, `.successfulPut,successfulScanCount`
`otel.metric.export.interval`	Opcional	Com que frequência, em milissegundos, coletar métricas do servidor HBase REST.	"60000"	Uma string especificando o número de milissegundos. Aceita somente números inteiros.

emr-hbase-thrift-server-propriedades métricas

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	
<custom_bean_name>	Opcional	N/D	O CloudWatch agente MBean desse agente deve coletar métricas, como`Hadoop:service=HBase,name=Thrift,sub=ThriftOne` . Você encontra exemplos de MBean nomes e suas métricas correspondentes nos exemplos de arquivos YAML do JMX para o Amazon EMR versão 7.0.	Uma string contendo a lista delimitada por vírgulas de métricas associadas ao. MBean Por exemplo, `.BatchGet_max,BatchGet_mean`	
`otel.metric.export.interval`	Opcional	Com que frequência, em milissegundos, coletar métricas	"60000"	Uma string especificando o número de milissegundos. Aceita somente	

Propriedade	Obrigatório	Descrição	Valor padrão	Possíveis valores	
		do servidor HBase Thrift.		números inteiros.	

Exemplos de configurações de métricas do sistema

O exemplo a seguir demonstra como configurar o CloudWatch agente para interromper a exportação de todas as métricas do sistema.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-system-metrics",
        "Properties": {},
        "Configurations": []
      }
    ]
  }
]
```

O exemplo a seguir configura o CloudWatch agente para exportar as métricas padrão do sistema. Isso é uma maneira rápida de fazer com que o agente volte a exportar apenas as métricas padrão do sistema se você já as tiver reconfigurado pelo menos uma vez. Essa redefinição também remove todas as métricas da aplicação que foram reconfiguradas anteriormente.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": []
  }
]
```

O exemplo a seguir configura o cluster para exportar as métricas `cpu`, `mem` e `disk`.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-system-metrics",
        "Properties": {
          "metrics_collection_interval": "20"
        },
        "Configurations": [
          {
            "Classification": "cpu",
            "Properties": {
              "metrics": "cpu_usage_guest,cpu_usage_idle",
              "metrics_collection_interval": "30",
              "drop_original_metrics": "cpu_usage_guest"
            }
          },
          {
            "Classification": "mem",
            "Properties": {
              "metrics": "mem_active"
            }
          },
           {
            "Classification": "disk",
            "Properties": {
              "metrics": "disk_used_percent",
              "resources": "/,/mnt",
              "drop_original_metrics": ""
            }
          }
        ]
      }
    ]
  }
]
```

A configuração do exemplo anterior tem as seguintes propriedades:

- A cada 30 segundos, o agente coleta a `cpu_guest` métrica para todos CPUs. Você pode encontrar a métrica agregada no CloudWatch `CWAgent > cluster.id, instance.id, node.type, service.name` namespace.
- A cada 30 segundos, o agente coleta a `cpu_idle` métrica para todos CPUs. Você pode encontrar a métrica agregada no CloudWatch `CWAgent > cluster.id, instance.id, node.type, service.name` namespace. O agente também coleta as métricas por CPU. Você pode encontrá-las no mesmo namespace. O agente coleta essa métrica porque a propriedade `drop_original_metrics` não contém `cpu_idle`, então o agente não ignora a métrica.
- A cada 20 segundos, o agente coleta a métrica `mem_active`. Você pode encontrar a métrica agregada no CloudWatch `CWAgent > cluster.id, instance.id, node.type, service.name` namespace.
- A cada 20 segundos, o agente coleta as métricas `disk_used_percent` das montagens `/` e `/mnt` do disco. Você pode encontrar as métricas agregadas no CloudWatch `CWAgent > cluster.id, instance.id, node.type, service.name` namespace. O agente também coleta as métricas por montagem. Você pode encontrá-las no mesmo namespace. O agente coleta essa métrica porque a propriedade `drop_original_metrics` não contém `disk_used_percent`, então o agente não ignora a métrica.

Exemplos de configurações de métricas de aplicações

O exemplo a seguir configura o CloudWatch agente para interromper a exportação de métricas para o serviço Namenode do Hadoop.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-hadoop-hdfs-namenode-metrics",
        "Properties": {},
        "Configurations": []
      }
    ]
  }
]
```

O exemplo a seguir configura um cluster para exportar métricas de aplicação do Hadoop.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-hadoop-hdfs-namenode-metrics",
        "Properties": {
          "Hadoop:service=NameNode,name=FSNamesystem": "BlockCapacity,CapacityUsedGB",
          "otel.metric.export.interval": "20000"
        },
        "Configurations": []
      },
       {
        "Classification": "emr-hadoop-hdfs-datanode-metrics",
        "Properties": {
          "Hadoop:service=DataNode,name=JvmMetrics": "MemNonHeapUsedM",
          "otel.metric.export.interval": "30000"
        },
        "Configurations": []
      },
       {
        "Classification": "emr-hadoop-yarn-resourcemanager-metrics",
        "Properties": {
          "Hadoop:service=ResourceManager,name=CapacitySchedulerMetrics":
 "AllocateNumOps,NodeUpdateNumOps"
        },
        "Configurations": []
      }
    ]
  }
]
```

O exemplo anterior tem as seguintes propriedades:

- A cada 20 segundos, o agente coleta as métricas `BlockCapacity` e `CapacityUsedGB` das instâncias que executam o serviço Namenode do Hadoop.
- A cada 30 segundos, o agente coleta as métricas `MemNonHeapUsedM` das instâncias que executam o serviço Datanode do Hadoop.
- A cada 30 segundos, o agente coleta as `NodeUpdateNumOps` métricas `AllocateNumOps` e das instâncias que executam o YARN do Hadoop. ResourceManaager

Exemplo do Amazon Managed Service para Prometheus

O exemplo a seguir demonstra como configurar o CloudWatch agente para exportar métricas ao Amazon Managed Service for Prometheus.

Se você está atualmente exportando métricas ao Amazon Managed Service para Prometheus e quiser reconfigurar as métricas para o cluster e continuar exportando métricas ao Amazon Managed Service para Prometheus, você deve incluir as propriedades `metrics_destination` e `prometheus_endpoint`.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {
      "metrics_destination": "prometheus",
      "prometheus_endpoint": "http://amp-workspace/api/v1/remote_write"
    },
    "Configurations": []
  }
]
```

Para usar o CloudWatch agente para onde exportar métricas CloudWatch, use o exemplo a seguir.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {
      "metrics_destination": "cloudwatch"
    },
    "Configurations": []
  }
]
```

> **Note**
>
> O CloudWatch agente tem um exportador do Prometheus que renomeia determinados atributos. Para os rótulos de métricas padrão, o Amazon Managed Service for Prometheus usa caracteres de sublinhado no lugar dos pontos da Amazon. CloudWatch Se você usar o Amazon Managed Grafana para visualizar as métricas padrão no Amazon Managed Service

para Prometheus, os rótulos aparecerão como `cluster_id`, `instance_id`, `node_type`, and `service_name`.

Configuração do CloudWatch agente para o Amazon EMR 7.0.0

Você pode configurar o CloudWatch agente da Amazon para usar métricas adicionais do sistema além daquelas fornecidas pela configuração padrão do CloudWatch agente. A configuração da versão 7.0.0 requer o uso de ações de bootstrap, das quais fornecemos exemplos nas seções a seguir. Em uma próxima versão, o Amazon EMR fornecerá opções adicionais de configuração por meio da API do Amazon EMR.

Tópicos

- Configuração de métricas adicionais do sistema com o Amazon EMR 7.0.0
- Configuração de métricas de aplicações com o Amazon EMR 7.0.0
- Configuração do Amazon Managed Service for Prometheus como armazenamento em nuvem para métricas com o Amazon EMR 7.0.0

Configuração de métricas adicionais do sistema com o Amazon EMR 7.0.0

Use as seguinte etapas para configurar o agente para usar um conjunto diferente de métricas do sistema no Amazon EMR 7.0.0:

1. Crie ou escolha um bucket na sua conta do Amazon S3 onde deseja armazenar os arquivos de configuração que especificam as métricas do CloudWatch agente.
2. Crie o arquivo de configuração `emr-amazon-cloudwatch-agent.json` com suas métricas preferenciais especificadas. Para isso, use um dos métodos explicados em Criação do arquivo de configuração do CloudWatch agente. Para obter mais informações sobre a estrutura do CloudWatch agente, consulte Criação ou edição manual do arquivo de configuração do CloudWatch agente no Guia CloudWatch do usuário da Amazon.
3. Em seguida, navegue até o aws-emr-utilities repositório em GitHub e baixe os scripts de métricas do sistema a seguir.

   - **`install_system_metrics_launcher.sh`**: um script que baixa e executa `install_system_metrics.sh` em segundo plano para que o nó possa concluir o bootstrapping.

- **install_system_metrics.sh**: um script que aguarda a instância em que é executado para concluir o bootstrapping e, em seguida, baixa e aplica a configuração no arquivo JSON.

4. Abra cada arquivo SH e substitua *amzn-s3-demo-bucket* pelo nome do seu bucket da Etapa 1.
5. Carregue um arquivo JSON e dois arquivos SH no seu bucket do S3.
6. Agora, você pode navegar até o console do Amazon EMR e criar um cluster com o CloudWatch agente. Em EMR, EC2 no painel de navegação à esquerda, selecione Clusters e, em seguida, Create cluster.
7. Na seção Nome e aplicações, escolha uma versão 7.0.0 ou superior do Amazon EMR.
8. Em Pacote de aplicações, selecione o pacote ou as aplicações que deseja instalar no cluster e inclua Agente do Amazon CloudWatch Agent nas seleções.
9. Na seção Ações de bootstrap, selecione Adicionar.
   - Em Nome, insira **install_system_metrics_launcher.sh**.
   - Na Localização do script, insira **s3://*amzn-s3-demo-bucket*/install_system_metrics_launcher.sh**. Substitua *amzn-s3-demo-bucket* pelo caminho para o bucket do S3.
   - Deixe o bloco Argumentos vazio.
10. Selecione Adicionar ação de bootstrap.
11. Continue criando o cluster para atender às necessidades da sua workload.

Quando seu cluster é executado, o CloudWatch agente publica as métricas do sistema que você especificou no arquivo de configuração para CloudWatch.

Configuração de métricas de aplicações com o Amazon EMR 7.0.0

Você pode configurar o CloudWatch agente da Amazon para publicar métricas de aplicações para o HDFS e o YARN, além das métricas do sistema. Use as etapas a seguir para configurar o agente para publicar as métricas de aplicações.

1. Crie ou escolha um bucket na sua conta do Amazon S3 onde deseja armazenar os arquivos de configuração que especificam as métricas do CloudWatch agente.
2. Em seguida, navegue até o aws-emr-utilities repositório em GitHub e baixe os scripts a seguir.

- **install_app_metrics_launcher.sh**: um script que baixa e executa `install_app_metrics.sh` em segundo plano para que o nó possa concluir o bootstrapping.
- **install_app_metrics.sh**: um script que aguarda a instância em que é executado para concluir o bootstrapping e, em seguida, baixa e aplica a configuração nos arquivos YAML que você baixará em uma próxima etapa.

3. Abra cada arquivo e substitua *amzn-s3-demo-bucket* pelo nome do seu bucket da Etapa 1.
4. Em seguida, baixe os arquivos de mapeamento YAML a seguir. Para obter informações sobre como esses arquivos YAML são estruturados, consulte `javaagent`no OpenTelemetry Instrumentation for Java GitHub repositório.

   - **datanode-metrics.yaml**: a configuração das métricas de Hadoop DataNode.
   - **namenode-metrics.yaml**: a configuração das métricas de Hadoop NameNode.
   - **nodemanager-metrics.yaml**: a configuração das métricas de Yarn NodeManager.
   - **resourcemanager-metrics.yaml**: a configuração das métricas de Yarn ResourceManager.

5. Carregue os dois arquivos SH e os quatro arquivos YAML no bucket do S3.
6. Agora, você pode navegar até o console do Amazon EMR e criar um cluster com o CloudWatch agente. Em EMR, EC2 no painel de navegação à esquerda, selecione Clusters e, em seguida, Create cluster.
7. Na seção Nome e aplicações, escolha uma versão 7.0.0 ou superior do Amazon EMR.
8. Em Pacote de aplicações, selecione o pacote ou o grupo personalizado de aplicações que deseja instalar no cluster e inclua CloudWatch Agente nas seleções.
9. Na seção Ações de bootstrap, selecione Adicionar.

   - Em Nome, insira **install_app_metrics_launcher.sh**.
   - Na Localização do script, insira **s3://*amzn-s3-demo-bucket*/install_app_metrics_launcher.sh**. Substitua *amzn-s3-demo-bucket* pelo caminho para o bucket do S3.
   - Deixe o bloco Argumentos vazio.

10. Selecione Adicionar ação de bootstrap.
11. Continue criando o cluster para atender às necessidades da sua workload.

Quando seu cluster é executado, o CloudWatch agente publica as métricas de aplicações que você especificou com as métricas do sistema. CloudWatch

Configuração do Amazon Managed Service for Prometheus como armazenamento em nuvem para métricas com o Amazon EMR 7.0.0

Você pode configurar o CloudWatch agente da Amazon para publicar métricas no Amazon Managed Service for Prometheus em vez de. CloudWatch

> **Note**
>
> Você pode publicar as métricas do CloudWatch agente da Amazon no Amazon Managed Service for Prometheus ou CloudWatch na Amazon, mas não nos dois serviços para o mesmo cluster.

Para configurar o agente para publicar métricas no Amazon Managed Service for Prometheus, você deve adicionar `aps:RemoteWrite` AWS Identity and Access Management a permissão (IAM) ao perfil de instância da EC2 Amazon EMR. O seguinte exemplo de política contém a permissão necessária:

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "aps:RemoteWrite",
      "Resource": "*"
    }
  ]
}
```

Use o CloudWatch agente em um cluster do EMR para publicar métricas no Amazon Managed Service for Prometheus

Depois que a política de serviço tiver as permissões corretas, use as etapas a seguir para executar um cluster que usa o CloudWatch agente para publicar métricas no Amazon Managed Service for Prometheus.

1. Use o AWS Management Console ou a AWS CLI para criar um espaço de trabalho do Amazon Managed Service for Prometheus. Para obter mais informações, consulte Criação de um espaço de trabalho no Guia do usuário do Amazon Managed Service for Prometheus.
2. Crie ou escolha um bucket na sua conta do Amazon S3 onde deseja armazenar os arquivos de execução que especificam o Amazon Managed Service for Prometheus como armazenamento em nuvem.
3. Em seguida, navegue até o aws-emr-utilities repositório em GitHub e baixe os scripts a seguir.

   - **add_prometheus_endpoint_launcher.sh**: um script que baixa e executa `add_prometheus_endpoint.sh` em segundo plano para que o nó possa concluir o bootstrapping.
   - **add_prometheus_endpoint.sh**— Um script que aguarda a instância em que é executado para concluir o bootstrapping e, em seguida, configura o CloudWatch agente para publicar no endpoint do Amazon Managed Service for Prometheus, que você fornece como argumento ao executar seu cluster.
4. Abra cada arquivo e substitua *amzn-s3-demo-bucket* pelo nome do seu bucket da Etapa 2.
5. Use a AWS CLI para criar um cluster do EMR com a ação de `add_prometheus_endpoint_launcher.sh` bootstrap. No comando a seguir, substitua *amzn-s3-demo-bucket* pelo bucket que contém a ação de bootstrap e substitua *managedpro-remote-write-workspace-url* pelo endpoint de gravação remota do seu espaço de trabalho do Amazon Managed Service for Prometheus. Certifique-se de especificar um rótulo de versão `emr-7.0.0` ou superior do Amazon EMR.

```
aws emr create-cluster --name managedpro-cluster \
    --release-label emr-7.0.0 \
    --applications Name=Hadoop Name=AmazonCloudWatchAgent \
    --ec2-attributes KeyName=myKey --instance-type m7g.2xlarge \
    --instance-count 3 --use-default-roles
   --bootstrap-actions Name='Add Prometheus Endpoint',Path=s3://amzn-s3-demo-
bucket/add_prometheus_endpoint_launcher.sh,Args='managedpro-remote-write-workspace-
url'
```

Quando seu cluster é executado, o CloudWatch agente publica as métricas coletadas no Amazon Managed Service for Prometheus.

Como usar o Amazon Managed Service for Prometheus como fonte de dados do Amazon Managed Grafana

Depois que o Amazon EMR publicar as métricas do cluster no Amazon Managed Service for Prometheus, você poderá usar as seguintes etapas para visualizar as métricas com o Amazon Managed Grafana:

1. Use a AWS Management Console para criar um espaço de trabalho e um usuário do Amazon Managed Grafana com as permissões apropriadas. Para obter mais informações, consulte Criação de um espaço de trabalho no Guia do usuário do Amazon Managed Grafana.
2. Adicione seu espaço de trabalho do Amazon Managed Service for Prometheus como fonte de dados no Amazon Managed Grafana. Para obter mais informações, consulte Use AWS data source configuration to add Amazon Managed Service for Prometheus as a data source no Guia do usuário do Amazon Managed Grafana.

> **Note**
>
> O CloudWatch agente tem um exportador do Prometheus que renomeia determinados atributos. Para os rótulos de métricas padrão, o Amazon Managed Service for Prometheus usa caracteres de sublinhado no lugar dos pontos da Amazon. CloudWatch Portanto, se você usar o Amazon Managed Grafana para visualizar as métricas padrão no Amazon Managed Service for Prometheus, os rótulos aparecerão como `jobflow_id`, `instance_id` e `service_name`.
>
> Além disso, qualquer métrica de aplicações que o CloudWatch agente publica no Amazon Managed Service for Prometheus usa o rótulo em vez de. `job` `service_name` No entanto, as métricas do sistema continuam usando o rótulo `service_name`.

Considerações e limitações

- O CloudWatch agente nativo da Amazon está disponível para clusters criados com as versões 7.0.0 e superiores do Amazon EMR.
- A configuração do CloudWatch agente no Amazon EMR 7.0.0 requer o uso de ações de bootstrap. Em uma próxima versão, o Amazon EMR fornecerá opções adicionais de configuração por meio da API do Amazon EMR.

- Não é possível instalar o CloudWatch agente do Amazon EMR se você já tiver implantado o CloudWatch agente por outro método, como Agent (SSM AWS Systems Manager Agent), na região em que seu cluster foi criado. Isso resultará no encerramento do cluster com erros.
- A CloudWatch GetMetricData API oferece suporte a até 500 métricas por solicitação. Se o seu cluster do Amazon EMR contiver mais de 250 nós em um grupo ou frota de instâncias, os gráficos correspondentes no painel CloudWatch incorporado do EMR aparecerão em branco com o erro Muitas métricas porque essas métricas exigem dois pontos de dados por métrica no painel de visão geral do cluster. No entanto, ao filtrar os painéis do Grupo de Instâncias Principais ou de Tarefas (ou frota), você poderá visualizar os gráficos de até 500 nós por grupo de instâncias ou frota. Isso ocorre porque eles não exigem dois pontos de dados por métrica. Além de 500 nós por grupo de instâncias ou frota, o erro Muitas métricas também ocorre para as métricas nesses painéis.

CloudWatch histórico de versões do agente

A tabela a seguir lista a versão AmazonCloudWatchAgent incluída em cada versão do Amazon EMR, junto com os componentes instalados com a aplicação. Para obter as versões dos componentes em cada versão, consulte a seção Versão de componentes da sua versão em Versões 7.x do Amazon EMR, Versões de lançamento 6.x do Amazon EMR ou Versões de lançamento 5.x do Amazon EMR.

AmazonCloudWatchAgent informações sobre a versão

Rótulo de versão do Amazon EMR	AmazonCloudWatchAgent Versão	Componentes instalados com o AmazonCloudWatchAgent
emzn-7.9.0	1.32.2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent , emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resour

Rótulo de versão do Amazon EMR	AmazonCloudWatchAgent Versão	Componentes instalados com o AmazonCloudWatchAgent
		cemanager, hadoop-yarn-timeline-server
emzn-7.8.0	1.32.2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.7.0	1.32.2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	AmazonCloudWatchAgent Versão	Componentes instalados com o AmazonCloudWatchAgent
emr-7.6.0	1.32.2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.5.0	1.32.2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	AmazonCloudWatchAgent Versão	Componentes instalados com o AmazonCloudWatchAgent
emr-7.4.0	1.32.2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.3.0	1.32.2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	AmazonCloudWatchAgent Versão	Componentes instalados com o AmazonCloudWatchAgent
emr-7.2.0	1.32.2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.1.0	1.32.2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.0.0	1.31.31.1	adot-java-agent, emr-amazon-cloudwatch-agent

Delta Lake

O Delta Lake é um framework de camada de armazenamento para arquiteturas de lakehouse comumente desenvolvidas no Amazon S3. Com as versões 6.9.0 e posteriores do Amazon EMR, é possível usar o Apache Spark 3.x em clusters do Amazon EMR com tabelas Delta Lake. Para obter mais informações sobre lakehouses com Delta Lake, consulte https://delta.io/.

A tabela a seguir lista a versão do Delta incluída na versão mais recente da série 7.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Delta.

Para a versão dos componentes instalados com o Delta nesta versão, consulte Versões de componentes da versão 7.9.0.

Informações sobre a versão Delta do emr-7.9.0

Rótulo de versão do Amazon EMR	Versão Delta	Componentes instalados com o Delta
emr-7.9.0	Delta 3.3.0-amzn-1	Not available.

A tabela a seguir lista a versão do Delta incluída na versão mais recente da série 6.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Delta.

Para obter a versão dos componentes instalados com o Delta nessa versão, consulte Release 6.15.0 Component Versions.

Informações de versão do Delta para o emr-6.15.0

Rótulo de versão do Amazon EMR	Versão Delta	Componentes instalados com o Delta
emr-6.15.0	Delta 2.4.0	Not available.

Introdução ao Delta Lake

O Delta Lake é um projeto de código aberto que ajuda a implementar arquiteturas modernas de data lake comumente desenvolvidas no Amazon S3. O Delta Lake oferece os seguintes recursos:

- Transações atômicas, consistentes, isoladas e duráveis (ACID) no Spark. Os leitores têm uma visualização consistente da tabela durante um trabalho do Spark.
- Tratamento escalável de metadados com processamento distribuído pelo Spark.
- Combina casos de uso de transmissão e em lote com a mesma tabela Delta.
- Aplicação automática de esquema para evitar registros incorretos durante a ingestão de dados.
- Viagem no tempo com versionamento de dados.
- Oferece suporte a operações de mesclagem, atualização e exclusão para casos de uso complexos, como captura de dados de alteração (CDC), atualizações de transmissão e muito mais.

Usar um cluster com o Delta Lake instalado

Tópicos

- Usar um cluster do Delta Lake com o Flink
- Usar um cluster Delta Lake com o Trino
- Usar um cluster do Delta Lake com o Spark
- Usar um cluster do Delta Lake com Spark e AWS Glue

Usar um cluster do Delta Lake com o Flink

Com o Amazon EMR versão 6.11 e versões posteriores, você pode usar o Delta Lake com seu cluster do Flink. Os exemplos a seguir usam o AWS CLI para trabalhar com o Delta Lake em um cluster Amazon EMR Flink.

> **Note**
>
> O Amazon EMR oferece suporte à DataStream API do Flink quando você usa o Delta Lake com um cluster do Flink.

Criar um cluster Delta Lake

1. Crie um arquivo `delta_configurations.json`, com o seguinte conteúdo:

```
[{"Classification":"delta-defaults",
```

```
    "Properties":{"delta.enabled":"true"}}]
```

2. Crie um cluster com a configuração a seguir. Substitua a `example Amazon S3 bucket path` e o `subnet ID` por seus próprios.

```
aws emr create-cluster
--release-label emr-6.11.0
--applications Name=Flink
--configurations file://delta_configurations.json
--region us-east-1  --name My_Spark_Delta_Cluster
--log-uri  s3://amzn-s3-demo-bucket/
--instance-type m5.xlarge
--instance-count 3
--service-role EMR_DefaultRole_V2
--ec2-attributes
 InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef0
```

Inicializar uma sessão do yarn do Flink

Para inicializar uma sessão do yarn do Flink, execute o seguinte comando:

```
flink-yarn-session -d
```

Compilar um trabalho no Flink com Delta Lake

Os exemplos a seguir mostram como usar o sbt ou o Maven para criar seu trabalho no Flink com o Delta Lake.

sbt

O sbt é uma ferramenta de build para Scala que pode ser usada com pouca ou nenhuma configuração quando você tem pequenos projetos.

```
libraryDependencies ++= Seq(
  "io.delta" %% "delta-flink" % deltaConnectorsVersion % "provided",
  "io.delta" %% "delta-standalone" % deltaConnectorsVersion % "provided",
  "org.apache.flink" %% "flink-clients" % flinkVersion % "provided",
  "org.apache.flink" %% "flink-parquet" % flinkVersion % "provided",
  "org.apache.hadoop" % "hadoop-client" % hadoopVersion % "provided",
  "org.apache.flink" % "flink-table-common" % flinkVersion % "provided",
```

```
  "org.apache.flink" %% "flink-table-runtime" % flinkVersion % "provided")
```

Maven

O Maven é uma ferramenta de automação de build de código aberto da Apache Software Foundation. Com o Maven, é possível compilar, publicar e implantar um trabalho do Flink com o Delta Lake no Amazon EMR.

```
<project>
<properties>
    <scala.main.version>2.12</scala.main.version>
    <delta-connectors-version>0.6.0</delta-connectors-version>
    <flink-version>1.16.1</flink-version>
    <hadoop-version>3.1.0</hadoop-version>
</properties>

<dependencies>
    <dependency>
        <groupId>io.delta</groupId>
        <artifactId>delta-flink</artifactId>
        <version>$delta-connectors-version</version>
        <scope>provided</scope>
    </dependency>
    <dependency>
        <groupId>io.delta</groupId>
        <artifactId>delta-standalone_$scala-main-version</artifactId>
        <version>$delta-connectors-version</version>
        <scope>provided</scope>
    </dependency>
    <dependency>
        <groupId>org.apache.flink</groupId>
        <artifactId>flink-clients</artifactId>
        <version>$flink-version</version>
        <scope>provided</scope>
    </dependency>
    <dependency>
        <groupId>org.apache.flink</groupId>
        <artifactId>flink-parquet</artifactId>
        <version>$flink-version</version>
        <scope>provided</scope>
    </dependency>
    <dependency>
        <groupId>org.apache.hadoop</groupId>
```

```
        <artifactId>hadoop-client</artifactId>
        <version>$hadoop-version</version>
        <scope>provided</scope>
    </dependency>
    <dependency>
        <groupId>org.apache.flink</groupId>
        <artifactId>flink-table-common</artifactId>
        <version>$flink-version</version>
        <scope>provided</scope>
    </dependency>
    <dependency>
        <groupId>org.apache.flink</groupId>
        <artifactId>flink-table-runtime</artifactId>
        <version>$flink-version</version>
        <scope>provided</scope>
    </dependency>
</dependencies>
```

Gravar em uma tabela Delta com a API Flink Datastream

Use o exemplo a seguir para criar um DeltaSink para gravar na tabela com um `deltaTablePath`:

```
public static DataStream<RowData> createDeltaSink(
        DataStream<RowData> stream,
        String deltaTablePath,
        RowType rowType) {
    Configuration configuration = new Configuration();
    DeltaSink<RowData> deltaSink = DeltaSink
            .forRowData(
                    new org.apache.flink.core.fs.Path(deltaTablePath),
                    configuration,
                    rowType)
            .build();
    stream.sinkTo(deltaSink);
    return stream;
}
```

Ler uma tabela Delta com a API Flink Datastream

Use o exemplo a seguir para criar um limite DeltaSource para leitura da tabela com um `deltaTablePath`:

```
public static DataStream<RowData> createBoundedDeltaSourceAllColumns(
        StreamExecutionEnvironment env,
        String deltaTablePath) {
    Configuration configuration = new Configuration();
    DeltaSource<RowData> deltaSource = DeltaSource
            .forBoundedRowData(
                    new org.apache.flink.core.fs.Path(deltaTablePath),
                    configuration)
            .build();

    return env.fromSource(deltaSource, WatermarkStrategy.noWatermarks(), "delta-
source");
}
```

Criação de coletores com suporte a vários clusters para Delta Lake independente

Use o exemplo a seguir para criar uma DeltaSink tabela para gravar com suporte a `deltaTablePath` e vários clusters:

```
public DataStream<RowData> createDeltaSink(
        DataStream<RowData> stream,
        String deltaTablePath) {
    Configuration configuration = new Configuration();
    configuration.set("spark.delta.logStore.s3.impl",
 "io.delta.storage.S3DynamoDBLogStore");
    configuration.set("spark.io.delta.storage.S3DynamoDBLogStore.ddb.tableName",
 "delta_log");
    configuration.set("spark.io.delta.storage.S3DynamoDBLogStore.ddb.region", "us-
east-1");

    DeltaSink<RowData> deltaSink = DeltaSink
        .forRowData(
            new Path(deltaTablePath),
            configuration,
            rowType)
        .build();
    stream.sinkTo(deltaSink);
    return stream;
}
```

Executar o trabalho do Flink

Use o seguinte comando para executar o trabalho:

```
flink run FlinkJob.jar
```

Usar um cluster Delta Lake com o Trino

Com as versões 6.9.0 e posteriores do Amazon EMR, você pode usar o Delta Lake com seu cluster do Trino.

Neste tutorial, usaremos o AWS CLI para trabalhar com o Delta Lake no cluster Trino do Amazon EMR.

Criar um cluster Delta Lake

1. Crie um arquivo, `delta_configurations.json`, e defina valores para o catálogo escolhido. Por exemplo, para usar o Hive Metastore como catálogo, o arquivo deve ter o seguinte conteúdo:

   ```
   [{"Classification":"delta-defaults",
       "Properties":{"delta.enabled":"true"}},
       {"Classification":"trino-connector-delta",
       "Properties":{"hive.metastore.uri":"thrift://localhost:9083"}}]
   ```

   Se você quiser usar o AWS Glue Catalog como sua loja, seu arquivo deve ter o seguinte conteúdo:

   ```
   [{"Classification":"delta-defaults",
       "Properties":{"delta.enabled":"true"}},
       {"Classification":"trino-connector-delta",
       "Properties":{"hive.metastore":"glue"}}]
   ```

2. Crie um cluster com a configuração a seguir, substituindo **`example Amazon S3 bucket path`** e **`subnet ID`** por suas informações.

   ```
   aws emr create-cluster
       --release-label emr-6.9.0
       --applications Name=Trino
       --configurations file://delta_configurations.json
   ```

```
    --region us-east-1  --name My_Spark_Delta_Cluster
    --log-uri  s3://amzn-s3-demo-bucket/
    --instance-type m5.xlarge
    --instance-count 2
    --service-role EMR_DefaultRole_V2
    --ec2-attributes
 InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef0
```

Inicializar a sessão do Trino para Delta Lake

Inicializar uma sessão do Trino, execute o comando a seguir

```
trino-cli --catalog delta
```

Gravar em uma tabela do Delta Lake

Crie e grave na tabela com os seguintes comandos SQL:

```
SHOW SCHEMAS;

CREATE TABLE default.delta_table (id  int, data varchar, category varchar) WITH
( location =  's3://amzn-s3-demo-bucket/<prefix>');

INSERT INTO default.delta_table VALUES  (1,'a','c1'), (2,'b','c2'), (3,'c','c3');
```

Ler uma tabela do Delta Lake

Leia a tabela com o seguinte comando SQL:

```
SELECT * from default.delta_table;
```

Usar um cluster do Delta Lake com o Spark

A partir da versão 6.9.0 do Amazon EMR, é possível usar o Delta Lake com seu cluster do Spark sem a necessidade de ações de bootstrap. Nas versões 6.8.0 e anteriores do Amazon EMR, é possível usar ações de bootstrap para pré-instalar as dependências necessárias.

Os exemplos a seguir usam o AWS CLI para trabalhar com o Delta Lake em um cluster do Amazon EMR Spark.

Para usar o Delta Lake no Amazon EMR com o AWS Command Line Interface, primeiro crie um cluster. Para obter informações sobre como especificar a classificação Delta Lake com AWS Command Line Interface, consulte Forneça uma configuração usando o AWS Command Line Interface ao criar um cluster ou Forneça uma configuração com o Java SDK ao criar um cluster.

1. Crie um arquivo `configurations.json`, com o seguinte conteúdo:

```
[{"Classification":"delta-defaults",  "Properties":{"delta.enabled":"true"} }]
```

2. Crie um cluster com a configuração a seguir, substituindo o **`bucket path`** do Amazon S3 de exemplo e a **`subnet ID`** por suas informações.

```
aws emr create-cluster
     --release-label  emr-6.9.0
     --applications Name=Spark
     --configurations file://delta_configurations.json
     --region us-east-1
     --name My_Spark_Delta_Cluster
     --log-uri  s3://amzn-s3-demo-bucket/
     --instance-type m5.xlarge
     --instance-count 2
     --service-role EMR_DefaultRole_V2
     --ec2-attributes
 InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef0
```

   Como alternativa, é possível criar um cluster do Amazon EMR e uma aplicação Spark com os seguintes arquivos como dependências JAR em um trabalho do Spark:

```
/usr/share/aws/delta/lib/delta-core.jar,
/usr/share/aws/delta/lib/delta-storage.jar,
/usr/share/aws/delta/lib/delta-storage-s3-dynamodb.jar
```

> **Note**
>
> Se você usa o Amazon EMR versões 7.0.0 ou superiores, use `/usr/share/aws/delta/lib/delta-spark.jar` em vez de. `/usr/share/aws/delta/lib/delta-core.jar`

Para obter mais informações, consulte Submitting Applications.

Para incluir uma dependência de jar no trabalho do Spark, é possível adicionar as seguintes propriedades de configuração à aplicação Spark:

```
--conf “spark.jars=/usr/share/aws/delta/lib/delta-core.jar,
     /usr/share/aws/delta/lib/delta-storage.jar,
     /usr/share/aws/delta/lib/delta-storage-s3-dynamodb.jar"
```

Para obter mais informações sobre as dependências de trabalhos do Spark, consulte Dependency Management.

Se você usa o Amazon EMR versões 7.0.0 ou superiores, adicione a configuração em vez disso. `/usr/share/aws/delta/lib/delta-spark.jar`

```
--conf “spark.jars=/usr/share/aws/delta/lib/delta-spark.jar,
     /usr/share/aws/delta/lib/delta-storage.jar,
     /usr/share/aws/delta/lib/delta-storage-s3-dynamodb.jar"
```

Inicializar uma sessão do Spark para Delta Lake

Os exemplos a seguir mostram como iniciar o shell interativo do Spark, usar o envio do Spark ou usar os Cadernos do Amazon EMR para trabalhar com o Delta Lake no Amazon EMR.

spark-shell

1. Conectar-se ao nó primário usando SSH. Para obter mais informações, consulte Connect to the primary node using SSH no Guia de gerenciamento do Amazon EMR.
2. Digite o seguinte comando para iniciar o shell do Spark. Para usar a PySpark concha, `spark-shell` substitua por`pyspark`.

```
spark-shell \
   --conf "spark.sql.extensions=io.delta.sql.DeltaSparkSessionExtension" \
   --conf
 "spark.sql.catalog.spark_catalog=org.apache.spark.sql.delta.catalog.DeltaCatalog"
```

Se você executa as versões 6.15.0 ou superiores do Amazon EMR, você também deve usar as configurações a seguir para utilizar o controle de acesso refinado baseado no Lake Formation com o Delta Lake.

```
spark-shell \
  --conf
 spark.sql.extensions=io.delta.sql.DeltaSparkSessionExtension,com.amazonaws.emr.recordserv
 \
  --conf
 spark.sql.catalog.spark_catalog=org.apache.spark.sql.delta.catalog.DeltaCatalog \

  --conf spark.sql.catalog.spark_catalog.lf.managed=true
```

spark-submit

1. Conectar-se ao nó primário usando SSH. Para obter mais informações, consulte Connect to the primary node using SSH no Guia de gerenciamento do Amazon EMR.
2. Insira o comando a seguir para iniciar a sessão do Spark no Delta Lake.

```
spark-submit
–conf "spark.sql.extensions=io.delta.sql.DeltaSparkSessionExtension"
–conf
 "spark.sql.catalog.spark_catalog=org.apache.spark.sql.delta.catalog.DeltaCatalog"
```

Se você executa as versões 6.15.0 ou superiores do Amazon EMR, você também deve usar as configurações a seguir para utilizar o controle de acesso refinado baseado no Lake Formation com o Delta Lake.

```
spark-submit \  `
--conf
 spark.sql.extensions=io.delta.sql.DeltaSparkSessionExtension,com.amazonaws.emr.recordserv
```

```
--conf
 spark.sql.catalog.spark_catalog=org.apache.spark.sql.delta.catalog.DeltaCatalog \

--conf spark.sql.catalog.spark_catalog.lf.managed=true
```

EMR Studio notebooks

Para inicializar uma sessão do Spark usando Cadernos do Amazon EMR Studio, configure a sessão do Spark usando o comando magic %%configure no Caderno do Amazon EMR, como no exemplo a seguir. Para obter mais informações, consulte Use EMR Notebooks magics no Guia de gerenciamento do Amazon EMR.

```
%%configure -f
{
  "conf": {
    "spark.sql.extensions":  "io.delta.sql.DeltaSparkSessionExtension",
     "spark.sql.catalog.spark_catalog":
 "org.apache.spark.sql.delta.catalog.DeltaCatalog"
  }
}
```

Se você executa as versões 6.15.0 ou superiores do Amazon EMR, você também deve usar as configurações a seguir para utilizar o controle de acesso refinado baseado no Lake Formation com o Delta Lake.

```
%%configure -f
{
  "conf": {
    "spark.sql.extensions":
 "io.delta.sql.DeltaSparkSessionExtension,com.amazonaws.emr.recordserver.connector.spark.sql
    "spark.sql.catalog.spark_catalog":
 "org.apache.spark.sql.delta.catalog.DeltaCatalog",
    "spark.sql.catalog.spark_catalog.lf.managed": "true"
  }
}
```

Gravar em uma tabela do Delta Lake

O exemplo a seguir mostra como criar um DataFrame e gravá-lo como um conjunto de dados do Delta Lake. O exemplo mostra como trabalhar com conjuntos de dados com o shell Spark enquanto estiver conectado ao nó primário usando SSH como usuário padrão do hadoop.

> **Note**
>
> Para colar exemplos de código no shell do Spark, digite :paste no prompt, cole o exemplo e pressione CTRL + D.

PySpark

O Spark inclui um shell baseado em Python, `pyspark`, que você pode usar para gerar protótipos de programas Spark gravados em Python. Assim como com `spark-shell`, invoque `pyspark` no nó primário.

```
## Create a DataFrame
data =  spark.createDataFrame([("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
("101",  "2015-01-01", "2015-01-01T12:14:58.597216Z"),
("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
("103",  "2015-01-01",  "2015-01-01T13:51:40.519832Z")],
["id", "creation_date",  "last_update_time"])

## Write a DataFrame as a Delta Lake dataset to the S3  location
spark.sql("""CREATE  TABLE IF NOT EXISTS delta_table (id string, creation_date
 string,
last_update_time string)
USING delta location
's3://amzn-s3-demo-bucket/example-prefix/db/delta_table'""");

data.writeTo("delta_table").append()
```

Scala

```
import org.apache.spark.sql.SaveMode
import org.apache.spark.sql.functions._

// Create a DataFrame
val data = Seq(("100",  "2015-01-01",  "2015-01-01T13:51:39.340396Z"),
("101",  "2015-01-01",  "2015-01-01T12:14:58.597216Z"),
```

```
("102",  "2015-01-01", "2015-01-01T13:51:40.417052Z"),
("103",  "2015-01-01",  "2015-01-01T13:51:40.519832Z")).toDF("id", "creation_date",
 "last_update_time")

// Write a DataFrame as a Delta Lake dataset to the S3  location
spark.sql("""CREATE  TABLE IF NOT EXISTS delta_table (id string,
creation_date string,
last_update_time string)
USING delta location
's3://amzn-s3-demo-bucket/example-prefix/db/delta_table'""");

data.write.format("delta").mode("append").saveAsTable("delta_table")
```

SQL

```
-- Create a Delta  Lake table with the S3 location
CREATE TABLE delta_table(id string,
creation_date string,
last_update_time string)
USING delta LOCATION
's3://amzn-s3-demo-bucket/example-prefix/db/delta_table';

-- insert data into the table
INSERT INTO delta_table VALUES  ("100", "2015-01-01",
 "2015-01-01T13:51:39.340396Z"),
("101",  "2015-01-01",  "2015-01-01T12:14:58.597216Z"),
("102",  "2015-01-01", "2015-01-01T13:51:40.417052Z"),
("103",  "2015-01-01", "2015-01-01T13:51:40.519832Z");
```

Ler uma tabela do Delta Lake

PySpark

```
ddf = spark.table("delta_table")
ddf.show()
```

Scala

```
val ddf =  spark.table("delta_table")
ddf.show()
```

SQL

```
SELECT * FROM delta_table;
```

Usar um cluster do Delta Lake com Spark e AWS Glue

Para usar o AWS Glue Catalog como Metastore para tabelas Delta Lake, crie um cluster com as etapas a seguir. Para obter informações sobre como especificar a classificação Delta Lake usando AWS Command Line Interface, consulte Forneça uma configuração usando o AWS Command Line Interface ao criar um cluster ou Forneça uma configuração usando o Java SDK ao criar um cluster.

Criar um cluster Delta Lake

1. Crie um arquivo `configurations.json`, com o seguinte conteúdo:

```
[{"Classification":"delta-defaults",
"Properties":{"delta.enabled":"true"}},
{"Classification":"spark-hive-site",
"Properties":
{"hive.metastore.client.factory.class":"com.amazonaws.glue.catalog.metastore.AWSGlueDataCat
```

2. Crie um cluster com a configuração a seguir, substituindo **`example Amazon S3 bucket path`** e **`subnet ID`** por suas informações.

```
aws emr create-cluster
    --release-label  emr-6.9.0
    --applications Name=Spark
    --configurations file://delta_configurations.json
    --region us-east-1
    --name My_Spark_Delta_Cluster
    --log-uri  s3://amzn-s3-demo-bucket/
    --instance-type m5.xlarge
    --instance-count 2
    --service-role EMR_DefaultRole_V2
    --ec2-attributes
 InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef0
```

Considerações e limitações

- O uso do Delta Lake é compatível com as versões 6.9.0 e posteriores do Amazon EMR. É possível usar o Apache Spark 3.x em clusters do Amazon EMR com tabelas Delta.
- Recomendamos usar o esquema de URI s3 para caminhos de localização do S3 em vez de s3a para melhorar a performance, a segurança e a confiabilidade. Para obter mais informações, consulte Working with storage and file systems.
- Com o Amazon EMR 7.0, não há Delta Universal Format (UniForm) suporte para convert-to-Iceberg declarações.
- Com o Amazon EMR 6.9 e 6.10, ao armazenar dados da tabela do Delta Lake no Amazon S3, os dados da coluna se tornam `NULL` após a operação de renomeação da coluna. Esse problema foi resolvido no Amazon EMR 6.11. Para obter mais informações sobre a operação experimental de renomeação de coluna, consulte Column rename operation no guia do usuário do Delta Lake.
- Se você criar um banco de dados no AWS Glue Data Catalog fora do Apache Spark, o banco de dados poderá ter um campo vazio`LOCATION`. Como o Spark não permite criar bancos de dados com uma propriedade de localização vazia, você receberá o seguinte erro ao usar o Spark no Amazon EMR para criar uma tabela Delta em um banco de dados do Glue e se o banco de dados tiver uma propriedade `LOCATION` vazia:

  ```
  IllegalArgumentException: Can not create a Path from an empty string
  ```

  Para resolver esse problema, crie o banco de dados no Catálogo de Dados com um caminho válido e não vazio para o campo `LOCATION`. Para ver as etapas para implementar a solução, consulte Exceção para argumento inválido ao criar uma tabela no Guia do usuário do Amazon Athena.

Histórico de versões do Delta

A tabela a seguir informa a versão do Delta incluída em cada versão do Amazon EMR, junto com os componentes instalados com a aplicação. Para obter as versões dos componentes em cada versão, consulte a seção Versão de componentes da sua versão em Versões 7.x do Amazon EMR, Versões de lançamento 6.x do Amazon EMR ou Versões de lançamento 5.x do Amazon EMR.

Informações da versão Delta

Rótulo de versão do Amazon EMR	Versão Delta	Componentes instalados com o Delta
emr-7.9.0	3.3.0-amzn-1	Not available.
emr-7.8.0	3.3.0-amzn-0	Not available.
emr-7.7.0	3.2.1-amzn-2	Not available.
emr-7.6.0	3.2.1-amzn-1	Not available.
emr-7.5.0	3.2.1-amzn-0	Not available.
emr-7.4.0	3.2.0-amzn-1	Not available.
emr-7.3.0	3.2.0-amzn-0	Not available.
emr-7.2.0	3.1.0	Not available.
emr-7.1.0	3.0.0	Not available.
emr-7.0.0	3.0.0	Not available.
emr-6.15.0	2.4.0	Not available.
emr-6.14.0	2.4.0	Not available.
emr-6.13.0	2.4.0	Not available.
emr-6.12.0	2.4.0	Not available.
emr-6.11.1	2.2.0	Not available.
emr-6.11.0	2.2.0	Not available.
emr-6.10.1	2.2.0	Not available.
emr-6.10.0	2.2.0	Not available.
emr-6.9.1	2.1.0	Not available.

Rótulo de versão do Amazon EMR	Versão Delta	Componentes instalados com o Delta
emr-6.9.0	2.1.0	Not available.

Apache Flink

O Apache Flink é um mecanismo de fluxo de dados de streaming que você pode usar para executar o processamento de streaming em tempo real em fontes de dados de alto throughput. O Flink suporta semântica de tempo de evento para out-of-order eventos, semântica de exatamente uma vez, controle de contrapressão e APIs otimizado para escrever aplicativos de streaming e em lote.

Além disso, o Flink tem conectores para fontes de dados de terceiros, como o seguinte:

- Amazon Kinesis Data Streams
- Apache Kafka
- Elasticsearch Connector do Flink
- API de streaming do Twitter
- Cassandra

O Amazon EMR oferece suporte ao Flink como um aplicativo do YARN para que você possa gerenciar recursos junto com outros aplicativos em um cluster. Flink-on-YARNpermite que você envie trabalhos temporários do Flink, ou você pode criar um cluster de longa execução que aceita vários trabalhos e aloca recursos de acordo com a reserva geral do YARN.

O Flink é incluído nas versões 5.1.0 e posteriores do Amazon EMR.

> **Note**
>
> O suporte para a classe `FlinkKinesisConsumer` foi adicionado no Amazon EMR versão 5.2.1.

A tabela a seguir lista a versão do Flink incluída na versão mais recente da série 7.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Flink.

Para a versão dos componentes instalados com o Flink nesta versão, consulte Versões de componentes da versão 7.9.0.

Informações sobre a versão do Flink para o emr-7.9.0

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-7.9.0	Flink 1.20.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta

A tabela a seguir lista a versão do Flink incluída na versão mais recente da série 6.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Flink.

Para obter a versão dos componentes instalados com o Flink nessa versão, consulte Release 6.15.0 Component Versions.

Informações de versão do Flink para o emr-6.15.0

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-6.15.0	Flink 1.17.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client,

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
		flink-jobmanager-config, hudi, delta-standalone-connectors

A tabela a seguir lista a versão do Flink incluída na versão mais recente da série 5.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Flink.

Para obter a versão dos componentes instalados com o Flink nesse lançamento, consulte Release 5.36.2 Component Versions.

Informações de versão do Flink para o emr-5.36.2

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.36.2	Flink 1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Tópicos

- Criar um cluster com o Flink
- Configurar o Flink no Amazon EMR
- Usar trabalhos do Flink no Amazon EMR
- Usar o shell Scala
- Localizar a interface Web do Flink
- Escalador automático do Flink

- Otimização dos tempos de reinício dos trabalhos para operações de recuperação e ajuste de escala de tarefas
- Usar trabalhos do Flink pelo Zeppelin no Amazon EMR
- Histórico de versões do Flink

Criar um cluster com o Flink

Você pode iniciar um cluster com o AWS Management Console, AWS CLI, ou um AWS SDK.

Iniciar um cluster com o Flink instalado no console

1. Abra o console do Amazon EMR em https://console.aws.amazon.com /emr.
2. Escolha Create cluster (Criar cluster), Go to advanced options (Ir para opções avançadas).
3. Para Software Configuration (Configuração do software), selecione EMR Release emr-5.1.0 ou posterior.
4. Selecione Flink como um aplicativo, juntamente com outros a serem instalados.
5. Selecione as outras opções conforme necessário e selecione Create cluster (Criar cluster).

Iniciar um cluster com o Flink na AWS CLI

- Crie o cluster usando o seguinte comando:

```
aws emr create-cluster --release-label emr-7.9.0 \
--applications Name=Flink \
--configurations file://./configurations.json \
--region us-east-1 \
--log-uri s3://myLogUri \
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole_V2 \
--ec2-attributes KeyName=MyKeyName,InstanceProfile=EMR_EC2_DefaultRole \
--steps Type=CUSTOM_JAR,Jar=command-runner.jar,Name=Flink_Long_Running_Session,\
Args=flink-yarn-session,-d
```

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

Configurar o Flink no Amazon EMR

Configurar o Flink com o Hive Metastore e o Catálogo do Glue

As versões 6.9.0 e posteriores do Amazon EMR oferecem suporte ao Hive Metastore e ao Catálogo do AWS Glue com o conector do Apache Flink para o Hive. Esta seção descreve as etapas necessárias para configurar o Catálogo do AWS Glue e o Hive Metastore com o Flink.

Tópicos

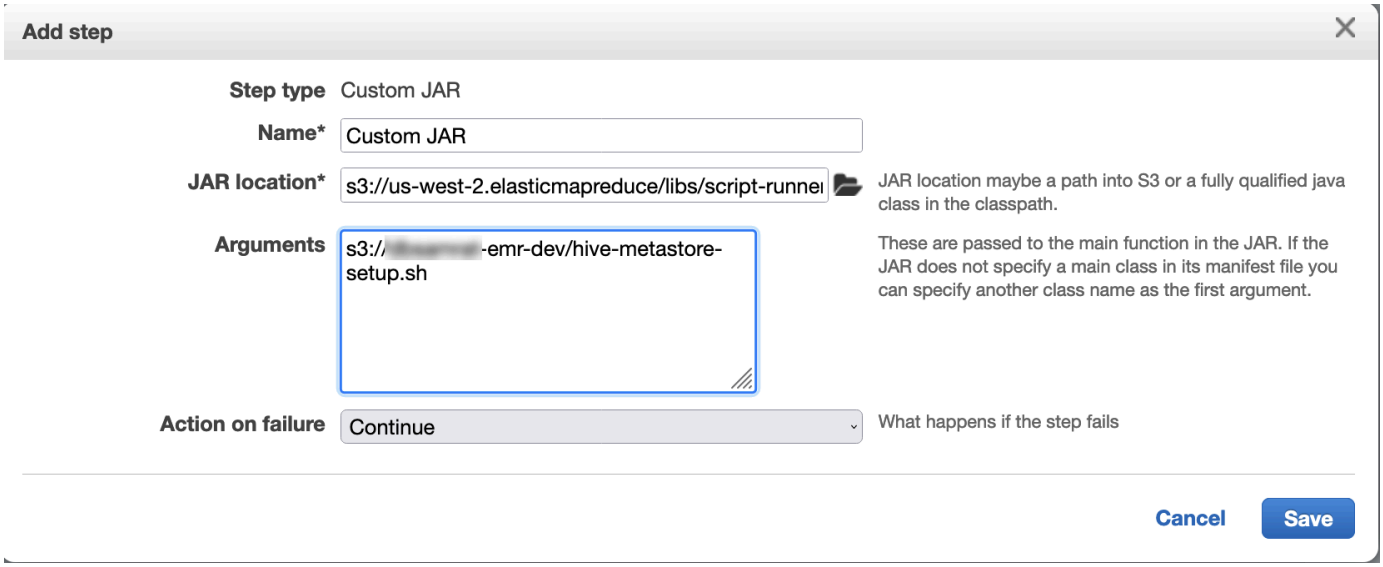
- Usar o Hive Metastore
- Usar o Catálogo de Dados do AWS Glue Data Catalog

Usar o Hive Metastore

1. Crie um cluster do EMR com a versão 6.9.0 ou posterior e pelo menos duas aplicações: Hive e Flink.
2. Use script runner para executar o script a seguir como função de etapa:

   `hive-metastore-setup.sh`

   ```
   sudo cp /usr/lib/hive/lib/antlr-runtime-3.5.2.jar /usr/lib/flink/lib
   sudo cp /usr/lib/hive/lib/hive-exec-3.1.3*.jar /lib/flink/lib
   sudo cp /usr/lib/hive/lib/libfb303-0.9.3.jar /lib/flink/lib
   sudo cp /usr/lib/flink/opt/flink-connector-hive_2.12-1.15.2.jar /lib/flink/lib
   sudo chmod 755 /usr/lib/flink/lib/antlr-runtime-3.5.2.jar
   sudo chmod 755 /usr/lib/flink/lib/hive-exec-3.1.3*.jar
   sudo chmod 755 /usr/lib/flink/lib/libfb303-0.9.3.jar
   sudo chmod 755 /usr/lib/flink/lib/flink-connector-hive_2.12-1.15.2.jar
   ```

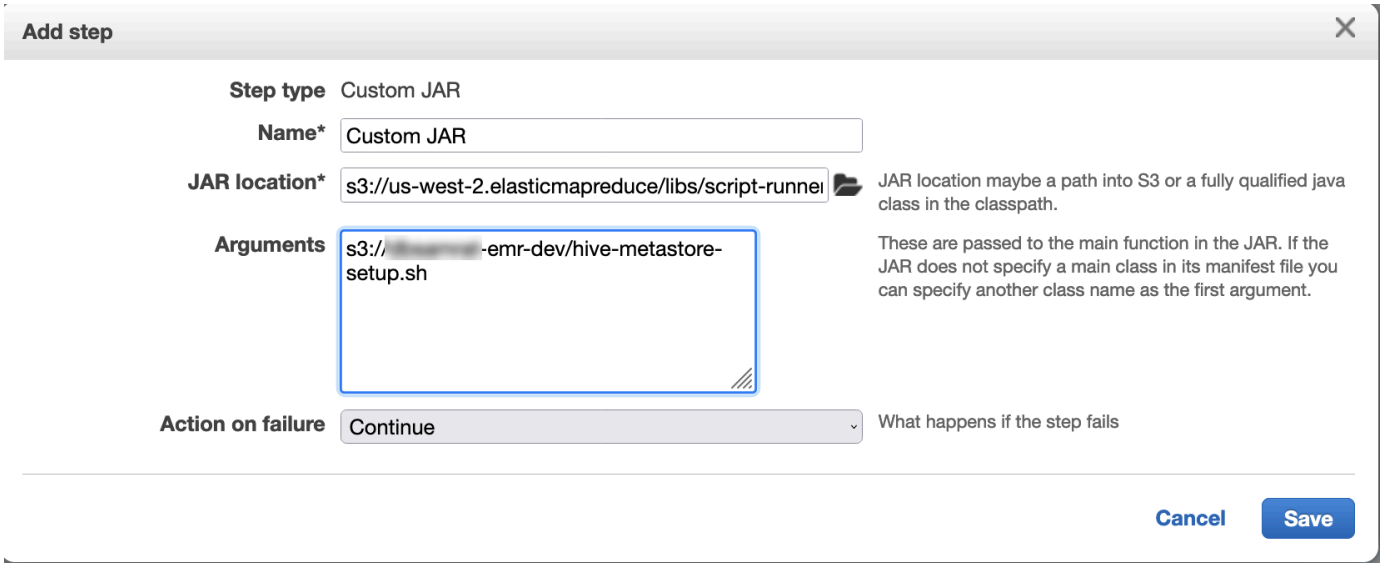


Usar o Catálogo de Dados do AWS Glue Data Catalog

1. Crie um cluster do EMR com a versão 6.9.0 ou posterior e pelo menos duas aplicações: Hive e Flink.
2. Selecione Usar com metadados da tabela do Hive nas configurações do Catálogo de Dados do AWS Glue para habilitar o Catálogo de Dados no cluster.
3. Use script runner para executar o seguinte script como função de etapa: Run commands and scripts on an Amazon EMR cluster:

   glue-catalog-setup.sh

   ```
   sudo cp /usr/lib/hive/auxlib/aws-glue-datacatalog-hive3-client.jar /usr/lib/flink/
   lib
   sudo cp /usr/lib/hive/lib/antlr-runtime-3.5.2.jar /usr/lib/flink/lib
   sudo cp /usr/lib/hive/lib/hive-exec-3.1.3*.jar /lib/flink/lib
   sudo cp /usr/lib/hive/lib/libfb303-0.9.3.jar /lib/flink/lib
   sudo cp /usr/lib/flink/opt/flink-connector-hive_2.12-1.15.2.jar /lib/flink/lib
   sudo chmod 755 /usr/lib/flink/lib/aws-glue-datacatalog-hive3-client.jar
   sudo chmod 755 /usr/lib/flink/lib/antlr-runtime-3.5.2.jar
   sudo chmod 755 /usr/lib/flink/lib/hive-exec-3.1.3*.jar
   sudo chmod 755 /usr/lib/flink/lib/libfb303-0.9.3.jar
   sudo chmod 755 /usr/lib/flink/lib/flink-connector-hive_2.12-1.15.2.jar
   ```



Configurar o Flink com um arquivo de configuração

É possível usar a API de configuração do Amazon EMR para configurar o Flink com um arquivo de configuração. Os arquivos que são configuráveis com a API são:

- `flink-conf.yaml`
- `log4j.properties`
- `flink-log4j-session`
- `log4j-cli.properties`

O principal arquivo de configuração para o Flink é `flink-conf.yaml`.

Configurar o número de slots de tarefa que são usados para o Flink na AWS CLI

1. Crie um arquivo `configurations.json`, com o seguinte conteúdo:

```
[
    {
      "Classification": "flink-conf",
      "Properties": {
        "taskmanager.numberOfTaskSlots":"2"
      }
    }
```

```
]
```

2. Em seguida, crie um cluster com a seguinte configuração:

```
aws emr create-cluster --release-label emr-7.9.0 \
--applications Name=Flink \
--configurations file://./configurations.json \
--region us-east-1 \
--log-uri s3://myLogUri \
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole_V2 \
--ec2-attributes KeyName=YourKeyName,InstanceProfile=EMR_EC2_DefaultRole
```

> **Note**
>
> Também é possível alterar algumas configurações com a API Flink. Para obter mais informações, consulte Concepts na documentação do Flink.
> Com as versões 5.21.0 e posteriores do Amazon EMR, você pode substituir as configurações de cluster e especificar classificações de configuração adicionais para cada grupo de instâncias em um cluster em execução. Você faz isso usando o console do Amazon EMR, o AWS Command Line Interface (AWS CLI) ou o AWS SDK. Para obter mais informações, consulte Supplying a Configuration for an Instance Group in a Running Cluster.

Opções de paralelismo

Como proprietário da aplicação, você sabe quais recursos atribuir a tarefas no Flink. Para os exemplos nesta documentação, use o mesmo número de tarefas que as instâncias de tarefa que você usa para a aplicação. Geralmente, recomendamos isso para o nível de paralelismo inicial, mas também é possível aumentar a granularidade do paralelismo usando slots de tarefa, que geralmente não excedem o número de núcleos virtuais por instância. Para obter mais informações sobre a arquitetura do Flink, consulte Concepts na documentação do Flink.

Configurar o Flink em um cluster do EMR com múltiplos nós primários

O Flink continua disponível durante o processo JobManager de failover do nó primário em um cluster do Amazon EMR com múltiplos nós primários. A partir do Amazon EMR 5.28.0, a JobManager alta disponibilidade também é habilitada automaticamente. Nenhuma configuração manual é necessária.

Com o Amazon EMR versões 5.27.0 ou anterior, JobManager é um único ponto de falha. Quando o JobManager falha, ele perde todos os estados de trabalho e não retoma os trabalhos em execução. Você pode ativar a JobManager alta disponibilidade configurando a contagem de tentativas do aplicativo, marcando o ponto de verificação e ativando o armazenamento ZooKeeper como estado para o Flink, conforme demonstra o exemplo a seguir:

```
[
  {
    "Classification": "yarn-site",
    "Properties": {
      "yarn.resourcemanager.am.max-attempts": "10"
    }
  },
  {
    "Classification": "flink-conf",
    "Properties": {
        "yarn.application-attempts": "10",
        "high-availability": "zookeeper",
        "high-availability.zookeeper.quorum": "%{hiera('hadoop::zk')}",
        "high-availability.storageDir": "hdfs:///user/flink/recovery",
        "high-availability.zookeeper.path.root": "/flink"
    }
  }
]
```

É necessário configurar o máximo de tentativas mestre do aplicativo para o YARN e as tentativas do aplicativo para o Flink. Para obter mais informações, consulte Configuration of YARN cluster high availability. Você também pode configurar o ponto de verificação do Flink para reiniciar a JobManager recuperação de trabalhos em execução de pontos de verificação concluídos anteriormente. Para obter mais informações, consulte Flink checkpointing.

Configurar o tamanho do processo de memória

Para versões do Amazon EMR que usam o Flink 1.11.x, é necessário configurar o tamanho total do processo de memória para () e JobManager (`jobmanager.memory.process.size`). TaskManager `taskmanager.memory.process.size` `flink-conf.yaml` É possível definir esses valores configurando o cluster com a API de configuração ou retirando manualmente os comentários desses campos via SSH. O Flink fornece os valores padrão a seguir.

- `jobmanager.memory.process.size`: 1600m

- `taskmanager.memory.process.size`: 1728m

Para excluir o metaspace e a sobrecarga da JVM, use o tamanho total da memória do Flink (`taskmanager.memory.flink.size`) em vez de `taskmanager.memory.process.size`. O valor padrão para `taskmanager.memory.process.size` é 1280m. Não é recomendável definir `taskmanager.memory.process.size` e `taskmanager.memory.process.size`.

Todas as versões do Amazon EMR que usam o Flink 1.12.0 e versões posteriores têm os valores padrão listados no conjunto de código aberto do Flink como valores padrão no Amazon EMR, então você não precisa configurá-los.

Configurar o tamanho do arquivo de saída de log

Os contêineres de aplicações Flink criam e gravam em três tipos de arquivos de log: arquivos `.out`, arquivos `.log` e arquivos `.err`. Somente os arquivos `.err` são compactados e removidos do sistema de arquivos, enquanto os arquivos de log `.log` e `.out` permanecem no sistema de arquivos. Para garantir que esses arquivos de saída continuem gerenciáveis e que o cluster continue estável, é possível configurar a alternância de logs `log4j.properties` para definir um número máximo de arquivos e limitar o tamanho deles.

Amazon EMR 5.30.0 e versões posteriores

A partir do Amazon EMR 5.30.0, o Flink usa a estrutura de registro em log log4j2 com o nome de classificação de configuração `flink-log4j.`. O exemplo de configuração a seguir demonstra o formato log4j2.

```
[
  {
    "Classification": "flink-log4j",
    "Properties": {
      "appender.main.name": "MainAppender",
      "appender.main.type": "RollingFile",
      "appender.main.append" : "false",
      "appender.main.fileName" : "${sys:log.file}",
      "appender.main.filePattern" : "${sys:log.file}.%i",
      "appender.main.layout.type" : "PatternLayout",
      "appender.main.layout.pattern" : "%d{yyyy-MM-dd HH:mm:ss,SSS} %-5p %-60c %x - %m
%n",
      "appender.main.policies.type" : "Policies",
      "appender.main.policies.size.type" : "SizeBasedTriggeringPolicy",
```

```
      "appender.main.policies.size.size" : "100MB",
      "appender.main.strategy.type" : "DefaultRolloverStrategy",
      "appender.main.strategy.max" : "10"
    },
  }
]
```

Amazon EMR 5.29.0 e versões anteriores

Com o Amazon EMR versão 5.29.0 e posteriores, o Flink usa o framework de registro em log log4j. O exemplo de configuração a seguir demonstra o formato log4j.

```
[
  {
    "Classification": "flink-log4j",
    "Properties": {
      "log4j.appender.file": "org.apache.log4j.RollingFileAppender",
      "log4j.appender.file.append":"true",
      # keep up to 4 files and each file size is limited to 100MB
      "log4j.appender.file.MaxFileSize":"100MB",
      "log4j.appender.file.MaxBackupIndex":4,
      "log4j.appender.file.layout":"org.apache.log4j.PatternLayout",
      "log4j.appender.file.layout.ConversionPattern":"%d{yyyy-MM-dd HH:mm:ss,SSS} %-5p
 %-60c %x - %m%n"
    },
  }
]
```

Configurar o Flink para ser executado com o Java 11

As versões 6.12.0 e posteriores do Amazon EMR oferecem suporte ao runtime do Java 11 para o Flink. As seções a seguir descrevem como configurar o cluster para fornecer suporte ao runtime do Java 11 para o Flink.

Tópicos

- Configurar o Flink para Java 11 ao criar um cluster
- Configurar o Flink para Java 11 em um cluster em execução
- Confirmar o runtime do Java para o Flink em um cluster em execução

Configurar o Flink para Java 11 ao criar um cluster

Realize as etapas a seguir para criar um cluster do EMR com o Flink e runtime do Java 11. O arquivo de configuração ao qual você adiciona suporte ao runtime do Java 11 é `flink-conf.yaml`.

Console

Para criar um cluster com o Flink e o runtime do Java 11 no console

1. Faça login no e abra AWS Management Console o console do Amazon EMR em https://console.aws.amazon.com /emr.
2. Escolha Clusters em EMR no painel de navegação e, EC2 em seguida, Criar cluster.
3. Selecione o Amazon EMR versão 6.12.0 ou posterior e escolha instalar a aplicação Flink. Selecione qualquer outra aplicação que você queira instalar no cluster.
4. Continue configurando o cluster. Na seção opcional Configurações de software, use a opção padrão Inserir configuração e insira a seguinte configuração:

```
[
    {
      "Classification": "flink-conf",
      "Properties": {
        "containerized.taskmanager.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
        "containerized.master.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
        "env.java.home":"/usr/lib/jvm/jre-11"
      }
    }
]
```

5. Continue configurando e iniciando o cluster.

AWS CLI

Criar um cluster com o Flink e o runtime do Java 11 na CLI

1. Crie um arquivo de configuração `configurations.json` que configure o Flink para usar o Java 11.

```
[
    {
      "Classification": "flink-conf",
```

```
      "Properties": {
        "containerized.taskmanager.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
        "containerized.master.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
        "env.java.home":"/usr/lib/jvm/jre-11"
      }
    }
]
```

2. Na AWS CLI, crie um novo cluster do EMR com o Amazon EMR versão 6.12.0 ou posterior e instale a aplicação Flink, conforme mostrado no exemplo a seguir:

```
aws emr create-cluster --release-label emr-6.12.0 \
--applications Name=Flink \
--configurations file://./configurations.json \
--region us-east-1 \
--log-uri s3://myLogUri \
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole_V2 \
--ec2-attributes KeyName=YourKeyName,InstanceProfile=EMR_EC2_DefaultRole
```

Configurar o Flink para Java 11 em um cluster em execução

Realize as etapas a seguir para atualizar um cluster do EMR em execução com o Flink e runtime do Java 11. O arquivo de configuração ao qual você adiciona suporte ao runtime do Java 11 é `flink-conf.yaml`.

Console

Para atualizar um cluster em execução com o Flink e o runtime do Java 11 no console

1. Faça login no e abra AWS Management Console o console do Amazon EMR em https://console.aws.amazon.com /emr.
2. Escolha Clusters em EMR EC2 no painel de navegação e, em seguida, selecione o cluster que você deseja atualizar.

> **Note**
>
> O cluster deve usar o Amazon EMR versão 6.12.0 ou posterior para oferecer suporte ao Java 11.

3. Selecione a guia Configuração.
4. Na seção Configurações do grupo de instâncias, selecione o grupo de instâncias Em execução que você deseja atualizar e escolha Reconfigurar no menu de ações da lista.
5. Reconfigure o grupo de instâncias com a opção Editar atributos, conforme mostrado a seguir. Selecione Adicionar nova configuração após cada.

Classificação	Propriedade	Valor
`flink-conf`	`containerized.taskmanager.env.JAVA_HOME`	`/usr/lib/jvm/jre-11`
`flink-conf`	`containerized.master.env.JAVA_HOME`	`/usr/lib/jvm/jre-11`
`flink-conf`	`env.java.home`	`/usr/lib/jvm/jre-11`

6. Selecione Salvar alterações para adicionar as configurações.

AWS CLI

Atualizar um cluster em execução para usar o Flink e o runtime do Java 11 na CLI

Você pode usar o comando `modify-instance-groups` para especificar configurações para cada grupo de instâncias em um cluster em execução.

1. Primeiro, crie um arquivo de configuração `configurations.json` que configure o Flink para usar o Java 11. No exemplo a seguir, *`ig-1xxxxxxx9`* substitua pelo ID do grupo de instâncias que você deseja reconfigurar. Salve o arquivo no mesmo diretório em que você executará o comando `modify-instance-groups`.

```
[
   {
      "InstanceGroupId":"ig-1xxxxxxx9",
      "Configurations":[
         {
            "Classification":"flink-conf",
            "Properties":{
              "containerized.taskmanager.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
              "containerized.master.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
              "env.java.home":"/usr/lib/jvm/jre-11"
            },
            "Configurations":[]
         }
      ]
   }
]
```

2. Na AWS CLI, execute o comando a seguir. Substitua o ID do grupo de instâncias que você deseja reconfigurar:

```
aws emr modify-instance-groups --cluster-id j-2AL4XXXXXX5T9 \
--instance-groups file://configurations.json
```

Confirmar o runtime do Java para o Flink em um cluster em execução

Para determinar o runtime do Java para um cluster em execução, faça login no nó primário com SSH, conforme descrito em Connect to the primary node with SSH. Em seguida, execute o seguinte comando:

```
ps -ef | grep flink
```

O comando `ps` com a opção `-ef` lista todos os processos que estão em execução no sistema. É possível filtrar essa saída com `grep` para encontrar menções à string `flink`. Revise a saída do valor do Ambiente de Execução Java (JRE), `jre-XX`. Na saída a seguir, `jre-11` indica que o Java 11 está selecionado em runtime para o Flink.

```
flink    19130     1  0 09:17 ?        00:00:15 /usr/lib/jvm/jre-11/bin/
java -Djava.io.tmpdir=/mnt/tmp -Dlog.file=/usr/lib/flink/log/flink-flink-
historyserver-0-ip-172-31-32-127.log -Dlog4j.configuration=file:/usr/lib/flink/conf/
```

```
log4j.properties -Dlog4j.configurationFile=file:/usr/lib/flink/conf/log4j.properties
 -Dlogback.configurationFile=file:/usr/lib/flink/conf/logback.xml -classpath /usr/lib/
flink/lib/flink-cep-1.17.0.jar:/usr/lib/flink/lib/flink-connector-files-1.17.0.jar:/
usr/lib/flink/lib/flink-csv-1.17.0.jar:/usr/lib/flink/lib/flink-json-1.17.0.jar:/usr/
lib/flink/lib/flink-scala_2.12-1.17.0.jar:/usr/lib/flink/lib/flink-table-api-java-
uber-1.17.0.jar:/usr/lib/flink/lib/flink-table-api-scala-bridge_2.12-1.17.0.
```

Como alternativa, faça login no nó primário com SSH e inicie uma sessão YARN do Flink com o comando `flink-yarn-session -d`. A saída mostra a Máquina Virtual Java (JVM) para Flink, `java-11-amazon-corretto` no exemplo a seguir:

```
2023-05-29 10:38:14,129 INFO  org.apache.flink.configuration.GlobalConfiguration
     [] - Loading configuration property: containerized.master.env.JAVA_HOME, /usr/lib/
jvm/java-11-amazon-corretto.x86_64
```

Usar trabalhos do Flink no Amazon EMR

Há várias maneiras de interagir com o Flink no Amazon EMR: por meio do console, da interface do Flink encontrada na interface de usuário de rastreamento e ResourceManager na linha de comando. Você pode enviar um arquivo JAR para uma aplicação Flink com qualquer uma destas opções. Depois de enviar um arquivo JAR, ele se torna um trabalho gerenciado pelo Flink JobManager. O JobManager está localizado no nó YARN que hospeda o daemon do Application Master da sessão Flink.

Você pode executar uma aplicação do Flink como um trabalho do YARN em um cluster de execução prolongada ou em um cluster transitório. Em um trabalho de execução prolongada, você pode enviar vários trabalhos do Flink para um cluster do Flink em execução no Amazon EMR. Se você executar um trabalho do Flink em um cluster transitório, seu cluster do Amazon EMR existirá somente pelo tempo necessário para executar a aplicação Flink; portanto, você pagará somente pelos recursos e pelo tempo usados. Você pode enviar um trabalho do Flink com a operação da API do Amazon `AddSteps` EMR, como um argumento de etapa para `RunJobFlow` a operação e por meio AWS CLI `add-steps` dos `create-cluster` comandos ou.

Iniciar uma aplicação YARN do Flink como uma etapa em um cluster de execução prolongada

Para iniciar uma aplicação Flink para a qual vários clientes possam enviar trabalhos por meio de operações de API do YARN, é necessário criar um cluster ou adicionar uma aplicação Flink a um

cluster já existente. Para obter instruções sobre como criar um novo cluster, consulte Criar um cluster com o Flink. Para iniciar uma sessão do YARN em um cluster atual, use as etapas a seguir no console, na AWS CLI ou no SDK do Java.

> **Note**
>
> O comando `flink-yarn-session` foi incluído no Amazon EMR versão 5.5.0 como um wrapper para o script `yarn-session.sh` para simplificar a execução. Se você usa uma versão anterior do Amazon EMR, substitua `bash -c "/usr/lib/flink/bin/yarn-session.sh -d"` por Arguments no console ou`Args` no comando AWS CLI .

Enviar um trabalho do Flink que está em um cluster existente no console

Envie a sessão do Flink usando o comando `flink-yarn-session` em um cluster já existente.

1. Abra o console do Amazon EMR em https://console.aws.amazon.com /emr.
2. Na lista de clusters, selecione o cluster que você iniciou anteriormente.
3. Na página de detalhes do cluster, selecione Steps (Etapas), Add Step (Adicionar etapa).
4. Use as diretrizes a seguir para inserir os parâmetros e escolha Adicionar.

Parameter	Descrição
Step type (Tipo de etapa)	JAR personalizado
Nome	Um nome para ajudar você a identificar a etapa. Por exemplo, *`<example-flink-step-name>`* .
Jar location (Local de jar)	command-runner.jar
Arguments (Argumentos)	O comando `flink-yarn-session` com argumentos adequados para seu aplicativo. Por exemplo, *`flink-yarn-session -d`* inicia uma sessão do Flink em seu cluster YARN em um estado desanexado (). -d Consulte a configuração do YARN na

Parameter	Descrição
	documentação do Flink mais recente para obter mais detalhes sobre argumentos.

Para enviar uma tarefa do Flink em um cluster existente com o AWS CLI

- Use o comando `add-steps` para adicionar um trabalho do Flink a um cluster de execução prolongada. O comando de exemplo a seguir especifica `Args="flink-yarn-session", "-d"` para iniciar uma sessão do Flink em seu cluster do YARN em um estado desconectado (`-d`). Consulte a configuração do YARN na documentação do Flink mais recente para obter mais detalhes sobre argumentos.

```
aws emr add-steps --cluster-id <j-XXXXXXXX> --steps Type=CUSTOM_JAR,Name=<example-flink-step-name>,Jar=command-runner.jar,Args="flink-yarn-session","-d"
```

Enviar o trabalho para uma aplicação Flink existente em um cluster de execução prolongada

Caso já tenha uma aplicação Flink em um cluster de execução prolongada, você poderá especificar o ID da aplicação Flink do cluster para enviar trabalhos para ele. Para obter o ID do aplicativo, execute `yarn application -list` na operação da YarnClientAPI AWS CLI ou por meio dela:

```
$ yarn application -list
16/09/07 19:32:13 INFO client.RMProxy: Connecting to ResourceManager at
 ip-10-181-83-19.ec2.internal/10.181.83.19:8032
Total number of applications (application-types: [] and states: [SUBMITTED, ACCEPTED,
 RUNNING]):1
Application-Id    Application-Name    Application-Type    User    Queue    State
 Final-State    Progress    Tracking-URL
application_1473169569237_0002    Flink session with 14 TaskManagers (detached)
   Apache Flink     hadoop    default            RUNNING          UNDEFINED
 100% http://ip-10-136-154-194.ec2.internal:33089
```

O ID do aplicativo para esta sessão do Flink é`application_1473169569237_0002`, que você pode usar para enviar trabalhos para o aplicativo a partir do AWS CLI ou de um SDK.

Example SDK para Java

```
List<StepConfig> stepConfigs = new ArrayList<StepConfig>();

HadoopJarStepConfig flinkWordCountConf = new HadoopJarStepConfig()
    .withJar("command-runner.jar")
    .withArgs("flink", "run", "-m", "yarn-cluster", "-yid",
 "application_1473169569237_0002", "-yn", "2", "/usr/lib/flink/examples/streaming/
WordCount.jar",
      "--input", "s3://amzn-s3-demo-bucket/pg11.txt", "--output", "s3://amzn-s3-demo-
bucket/alice2/");

StepConfig flinkRunWordCount = new StepConfig()
  .withName("Flink add a wordcount step")
  .withActionOnFailure("CONTINUE")
  .withHadoopJarStep(flinkWordCountConf);

stepConfigs.add(flinkRunWordCount);

AddJobFlowStepsResult res = emr.addJobFlowSteps(new AddJobFlowStepsRequest()
   .withJobFlowId("myClusterId")
   .withSteps(stepConfigs));
```

Example AWS CLI

```
aws emr add-steps --cluster-id <j-XXXXXXXX> \
--steps Type=CUSTOM_JAR,Name=Flink_Submit_To_Long_Running,Jar=command-runner.jar,\
Args="flink","run","-m","yarn-cluster","-yid","application_1473169569237_0002",\
"/usr/lib/flink/examples/streaming/WordCount.jar",\
"--input","s3://amzn-s3-demo-bucket/pg11.txt","--output","s3://amzn-s3-demo-bucket/
alice2/" \
--region <region-code>
```

Enviar um trabalho transitório do Flink

Os exemplos a seguir iniciam um cluster transitório que executa um trabalho do Flink e é terminado na conclusão.

Example SDK para Java

```
import java.util.ArrayList;
import java.util.List;
```

```
import com.amazonaws.AmazonClientException;
import com.amazonaws.auth.AWSCredentials;
import com.amazonaws.auth.AWSStaticCredentialsProvider;
import com.amazonaws.auth.profile.ProfileCredentialsProvider;
import com.amazonaws.services.elasticmapreduce.AmazonElasticMapReduce;
import com.amazonaws.services.elasticmapreduce.AmazonElasticMapReduceClientBuilder;
import com.amazonaws.services.elasticmapreduce.model.*;

public class Main_test {

 public static void main(String[] args) {
  AWSCredentials credentials_profile = null;
  try {
   credentials_profile = new ProfileCredentialsProvider("default").getCredentials();
  } catch (Exception e) {
   throw new AmazonClientException(
     "Cannot load credentials from .aws/credentials file. " +
       "Make sure that the credentials file exists and the profile name is specified
 within it.",
     e);
  }

  AmazonElasticMapReduce emr = AmazonElasticMapReduceClientBuilder.standard()
    .withCredentials(new AWSStaticCredentialsProvider(credentials_profile))
    .withRegion(Regions.US_WEST_1)
    .build();

  List<StepConfig> stepConfigs = new ArrayList<StepConfig>();
  HadoopJarStepConfig flinkWordCountConf = new HadoopJarStepConfig()
    .withJar("command-runner.jar")
    .withArgs("bash", "-c", "flink", "run", "-m", "yarn-cluster", "-yn", "2",
      "/usr/lib/flink/examples/streaming/WordCount.jar", "--input", "s3://path/to/
input-file.txt", "--output",
      "s3://path/to/output/");

  StepConfig flinkRunWordCountStep = new StepConfig()
    .withName("Flink add a wordcount step and terminate")
    .withActionOnFailure("CONTINUE")
    .withHadoopJarStep(flinkWordCountConf);

  stepConfigs.add(flinkRunWordCountStep);

  Application flink = new Application().withName("Flink");
```

```
  RunJobFlowRequest request = new RunJobFlowRequest()
    .withName("flink-transient")
    .withReleaseLabel("emr-5.20.0")
    .withApplications(flink)
    .withServiceRole("EMR_DefaultRole")
    .withJobFlowRole("EMR_EC2_DefaultRole")
    .withLogUri("s3://path/to/my/logfiles")
    .withInstances(new JobFlowInstancesConfig()
      .withEc2KeyName("myEc2Key")
      .withEc2SubnetId("subnet-12ab3c45")
      .withInstanceCount(3)
      .withKeepJobFlowAliveWhenNoSteps(false)
      .withMasterInstanceType("m4.large")
      .withSlaveInstanceType("m4.large"))
    .withSteps(stepConfigs);

  RunJobFlowResult result = emr.runJobFlow(request);
  System.out.println("The cluster ID is " + result.toString());

 }

}
```

Example AWS CLI

Use o subcomando `create-cluster` para criar um cluster transitório que termina quando o trabalho do Flink é concluído:

```
aws emr create-cluster --release-label emr-5.2.1 \
--name "Flink_Transient" \
--applications Name=Flink \
--configurations file://./configurations.json \
--region us-east-1 \
--log-uri s3://myLogUri \
--auto-terminate
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole_V2 \
--ec2-attributes KeyName=<YourKeyName>,InstanceProfile=EMR_EC2_DefaultRole \
--steps Type=CUSTOM_JAR,Jar=command-runner.jar,Name=Flink_Long_Running_Session,\
Args="bash","-c","\"flink run -m yarn-cluster /usr/lib/flink/examples/streaming/
WordCount.jar
--input s3://amzn-s3-demo-bucket/pg11.txt --output s3://amzn-s3-demo-bucket/alice/""
```

Usar o shell Scala

O shell Scala do Flink para clusters do EMR está configurado somente para iniciar novas sessões do YARN. Você pode usar o shell Scala seguindo o procedimento abaixo.

Usar o shell Flink Scala no nó primário

1. Faça login no nó primário com SSH, conforme descrito em Connect to the primary node with SSH.
2. Digite o seguinte para iniciar um shell:

   No Amazon EMR versão 5.5.0 e posterior, você pode usar o comando a seguir para iniciar um cluster do Yarn para o Scala Shell com um. TaskManager

   ```
   % flink-scala-shell yarn 1
   ```

   Em versões anteriores do Amazon EMR, use:

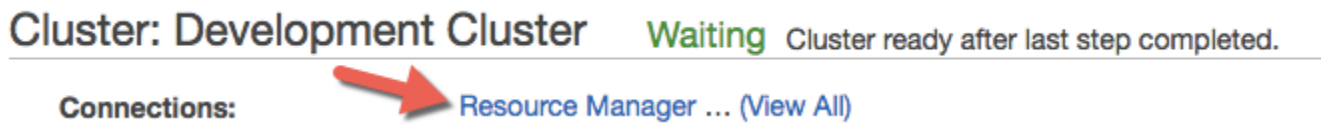
   ```
   % /usr/lib/flink/bin/start-scala-shell.sh yarn 1
   ```

   Isso inicia o shell Scala do Flink para que você possa usar o Flink interativamente. Assim como com outras interfaces e opções, você pode escalar o valor da opção `-n` usado no exemplo com base no número de tarefas que você deseja executar a partir do shell.

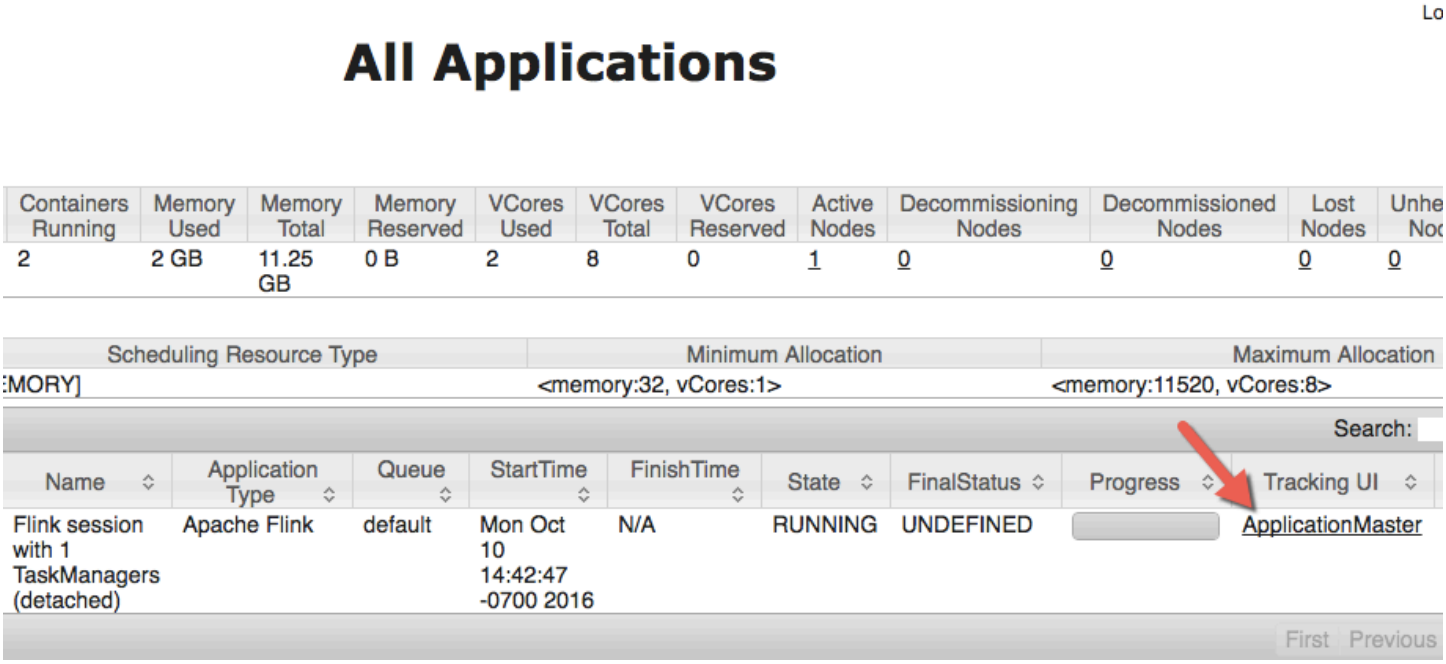
   Para obter mais informações, consulte Scala REPL na documentação oficial do Apache Flink.

Localizar a interface Web do Flink

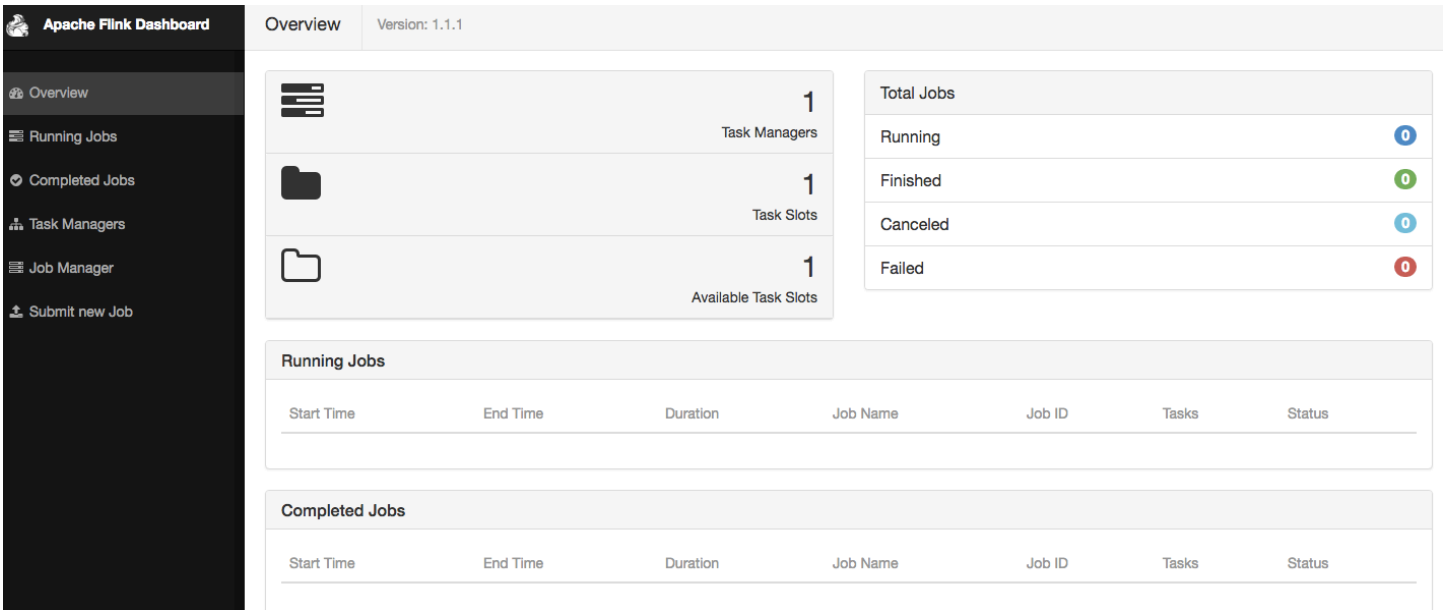
O Application Master que pertence à aplicação Flink hospeda a interface da Web do Flink. É uma forma alternativa de enviar um JAR como trabalho ou de visualizar o status atual de outros trabalhos. A interface da web do Flink ficará ativa desde que você tenha um sessão do Flink em execução. Se você já tiver um trabalho do YARN de longa duração ativo, siga as instruções no tópico Conecte-se ao nó primário com SSH no Guia de Gerenciamento do Amazon EMR para se conectar ao YARN. ResourceManager Por exemplo, se você configurou um túnel SSH e ativou um proxy em seu navegador, escolha a ResourceManager conexão em Conexões na página de detalhes do cluster do EMR.



Depois de encontrar o ResourceManager, selecione o aplicativo YARN que está hospedando uma sessão do Flink. Selecione o link na coluna Tracking UI.



Na interface da web do Flink, você pode visualizar a configuração, enviar seu próprio JAR personalizado como um trabalho ou monitorar os trabalhos em andamento.



Escalador automático do Flink

Visão geral

As versões 6.15.0 e superiores do Amazon EMR oferecem suporte ao escalador automático do Flink. A funcionalidade de escalador automático de trabalhos coleta métricas da execução de trabalhos de streaming do Flink e escala automaticamente os vértices individuais do trabalho. Isso reduz a contrapressão e satisfaz a meta de utilização definida por você.

Para obter mais informações, consulte a seção Autoscaler da documentação do Apache Flink Kubernetes Operator.

Considerações

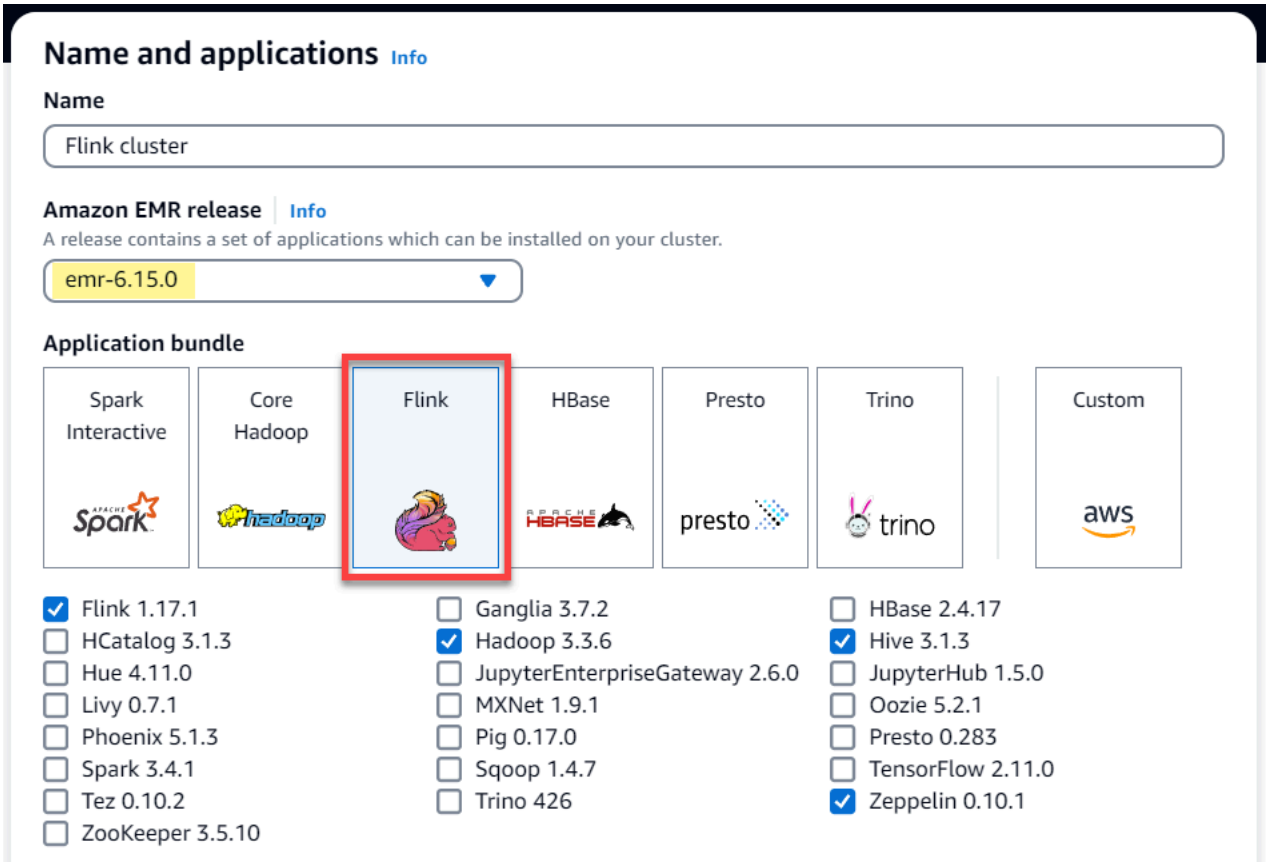
- O escalador automático do Flink é compatível com o Amazon EMR 6.15.0 e versões superiores.
- O escalador automático do Flink é compatível somente com trabalhos de streaming.
- Só há suporte para o agendador adaptável. Não há suporte para o agendador padrão.
- Recomendamos habilitar o ajuste de escala de clusters para permitir o provisionamento dinâmico de recursos. O Ajuste de Escala Gerenciado do Amazon EMR é preferencial porque a avaliação da métrica ocorre a cada cinco a dez segundos. Nesse intervalo, seu cluster pode se ajustar mais prontamente à mudança nos recursos necessários do cluster.

Habilitação do escalador automático

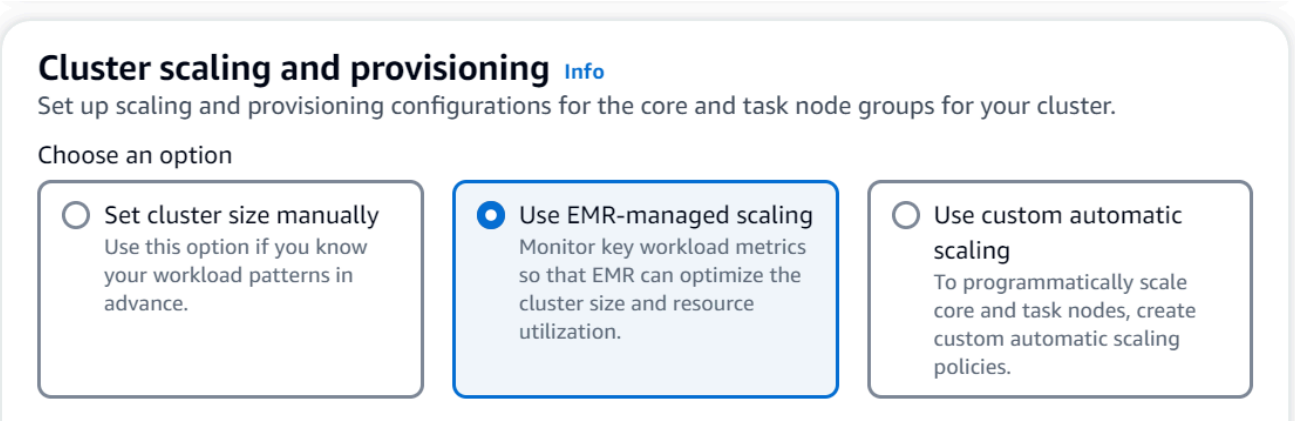
Use as etapas a seguir para habilitar o autoescalador Flink ao criar um Amazon EMR no cluster. EC2

1. No console do Amazon EMR, crie um cluster do EMR.

    a. Escolha a versão `emr-6.15.0` do Amazon EMR ou posterior. Selecione o pacote de aplicações Flink e quaisquer outras aplicações que você queira incluir no cluster.



b. Na opção Ajuste de escala e provisionamento de clusters, selecione Usar Ajuste de Escala Gerenciado do EMR.



2. Na seção Configurações de software, insira a configuração a seguir para habilitar o escalador automático do Flink. Em cenários de teste, defina o intervalo de decisão, o intervalo da janela de métricas e o intervalo de estabilização para um valor menor de modo que o trabalho tome imediatamente uma decisão de ajuste de escala para facilitar a verificação.

```
[
  {
```

```
    "Classification": "flink-conf",
    "Properties": {
      "job.autoscaler.enabled": "true",
      "jobmanager.scheduler": "adaptive",
      "job.autoscaler.stabilization.interval": "60s",
      "job.autoscaler.metrics.window": "60s",
      "job.autoscaler.decision.interval": "10s",
      "job.autoscaler.debug.logs.interval": "60s"
    }
  }
]
```

3. Selecione ou defina quaisquer outras configurações da maneira que preferir e crie o cluster habilitado pelo escalador automático do Flink.

Configurações do escalador automático

Esta seção aborda a maioria das configurações que você pode alterar com base nas suas necessidades específicas.

> **Note**
>
> Com configurações baseadas em tempo, como `time`, `interval` e `window`, a unidade padrão quando nenhuma unidade é especificada é milissegundos. Portanto, um valor de `30` sem sufixo é igual a 30 milissegundos. Para outras unidades de tempo, inclua o sufixo apropriado de `s` para segundos, `m` para minutos ou `h` para horas.

Tópicos

- Configurações de loop do escalador automático
- Agregação de métricas e configurações de histórico
- Configurações no nível do vértice do trabalho
- Configurações de processamento do backlog
- Configurações da operação de ajuste de escala

Configurações de loop do escalador automático

O escalador automático busca as métricas do nível do vértice do trabalho para cada intervalo de tempo configurável, converte-as em ações de escala, estima o novo paralelismo do vértice do trabalho e o recomenda ao agendador de trabalhos. As métricas são coletadas somente após o horário de reinício do trabalho e o intervalo de estabilização do cluster.

Chave de configuração	Valor padrão	Descrição	Exemplos de valores
`job.autoscaler.enabled`	`false`	Habilita o ajuste de escala automático no cluster do Flink.	`true`, `false`
`job.autoscaler.decision.interval`	`60s`	Intervalo de decisões do escalador automático.	`30` (a unidade padrão é milissegundos), `5m`, `1h`
`job.autoscaler.restart.time`	`3m`	O tempo de reinício esperado a ser usado até que o operador possa determiná-lo de forma confiável com o histórico.	`30` (a unidade padrão é milissegundos), `5m`, `1h`
`job.autoscaler.stabilization.interval`	`300s`	O período de estabilização no qual nenhum novo ajuste de escala será executado.	`30` (a unidade padrão é milissegundos), `5m`, `1h`
`job.autoscaler.debug.logs.interval`	`300s`	Intervalo dos logs de depuração do escalador automático.	`30` (a unidade padrão é milissegundos), `5m`, `1h`

Agregação de métricas e configurações de histórico

O escalador automático busca as métricas, faz a agregação delas ao longo de uma janela deslizante com base no tempo e elas são avaliadas em decisões de ajuste de escala. O histórico de decisões

do ajuste de escala para cada vértice do trabalho é utilizado para estimar novos paralelismos. Eles têm validade baseada na idade e no tamanho do histórico (pelo menos 1).

Chave de configuração	Valor padrão	Descrição	Exemplos de valores
`job.autoscaler.metrics.window`	`600s`	`Scaling metrics aggregation window size.`	`30` (a unidade padrão é milissegundos), `5m`, `1h`
`job.autoscaler.history.max.count`	`3`	Número máximo de decisões de ajuste de escala anteriores a serem retidas por vértice.	`1` para `Integer.MAX_VALUE`
`job.autoscaler.history.max.age`	`24h`	Número mínimo de decisões de ajuste de escala anteriores a serem retidas por vértice.	`30` (a unidade padrão é milissegundos), `5m`, `1h`

Configurações no nível do vértice do trabalho

O paralelismo de cada vértice do trabalho é modificado com base na utilização de destino e vinculado aos limites mínimo e máximo do paralelismo. Não é recomendável definir uma utilização de destino próxima de 100% (ou seja, valor de 1), e o limite de utilização funciona como um buffer para lidar com as flutuações intermediárias da carga.

Chave de configuração	Valor padrão	Descrição	Exemplos de valores
`job.autoscaler.target.utilization`	`0.7`	Utilização de destino do vértice.	`0 - 1`
`job.autoscaler.target.utilization.boundary`	`0.4`	Limite de utilização de destino do vértice. O ajuste de escala não será realizado se a taxa de	`0 - 1`

Chave de configuração	Valor padrão	Descrição	Exemplos de valores
		processamento atual estiver entre [target_rate / (target_utilizatio n - boundary) e (target_rate / (target_utilization + boundary)] .	
job.autoscaler.ver tex.min-parallelism	1	O paralelismo mínimo que o escalador automático pode usar.	0 - 200
job.autoscaler.ver tex.max-parallelism	200	O paralelismo máximo que o escalador automático pode usar. Observe que esse limite será ignorado se ele for maior que o paralelismo máximo definido na configuração do Flink ou diretamente em cada operador.	0 - 200

Configurações de processamento do backlog

O vértice do trabalho precisa de recursos extras para lidar com os eventos pendentes, ou backlogs, que se acumulam durante o período da operação de escala. Isso também é conhecido como duração de `catch-up`. Se o tempo de processamento do backlog exceder o valor de `lag - threshold` configurado, a utilização de destino do vértice do trabalho aumentará para o nível máximo. Isso ajuda a evitar operações desnecessárias de ajuste de escala durante o processamento do backlog.

Chave de configuração	Valor padrão	Descrição	Exemplos de valores
`job.autoscaler.backlog-processing.lag-threshold`	`5m`	Limite de atraso que evitará ajustes de escala desnecessários e removerá as mensagens pendentes responsáveis pelo atraso.	`30` (a unidade padrão é milissegundos), `5m`, `1h`
`job.autoscaler.catch-up.duration`	`15m`	A duração de destino para o processamento completo de qualquer backlog após uma operação de ajuste de escala. Defina como 0 para desabilitar o ajuste de escala baseado em backlog.	`30` (a unidade padrão é milissegundos), `5m`, `1h`

Configurações da operação de ajuste de escala

O escalador automático não executa a operação de redução vertical de escala imediatamente após uma operação de aumento vertical de escala dentro do período de carência. Isso evita ciclos desnecessários de up-down-up-down operações de balança causados por flutuações temporárias de carga.

Podemos usar a taxa de operação de redução vertical de escala para diminuir gradualmente o paralelismo e liberar recursos que atendam a picos de carga temporários. Isso também ajuda a evitar operações secundárias desnecessárias de aumento vertical de escala após uma grande operação de redução vertical de escala.

Podemos detectar uma operação ineficaz de aumento vertical de escala com base no histórico anterior de decisões de ajuste de escala de vértices de trabalho para evitar mais alterações no paralelismo.

Chave de configuração	Valor padrão	Descrição	Exemplos de valores
`job.autoscaler.scale-up.grace-period`	`1h`	Duração em que nenhuma redução vertical de escala de um vértice é permitida após ele ter sido aumentado.	`30` (a unidade padrão é milissegundos), `5m`, `1h`
`job.autoscaler.scale-down.max-factor`	`0.6`	Fator máximo de redução vertical de escala. Um valor de `1` significa que não há limite na redução vertical de escala; `0.6` significa que o trabalho só pode ser reduzido verticalmente com 60% do paralelismo original.	`0` - `1`
`job.autoscaler.scale-up.max-factor`	`100000.`	Taxa máxima de aumento vertical de escala. Um valor de `2.0` significa que o trabalho só pode aumentar a escala verticalmente com 200% do paralelismo atual.	`0` - `Integer.MAX_VALUE`
`job.autoscaler.scaling.effectiveness.detection.enabled`	`false`	Se deve habilitar a detecção de operações ineficazes de ajuste de escala e permitir que o escalador automático bloqueie novos aumentos verticais.	`true`, `false`

Otimização dos tempos de reinício dos trabalhos para operações de recuperação e ajuste de escala de tarefas

Quando uma tarefa falha ou ocorre uma operação de ajuste de escala, o Flink tenta reexecutar a tarefa com base no último ponto de verificação concluído. O processo de reinicialização pode levar

um minuto ou mais para ser executado, dependendo do tamanho do estado do ponto de verificação e do número de tarefas paralelas. Durante o período de reinicialização, as tarefas de backlog podem se acumular para o trabalho. No entanto, existem algumas maneiras de o Flink otimizar a velocidade de recuperação e reinicialização dos gráficos de execução para melhorar a estabilidade do trabalho.

Esta página descreve algumas maneiras que o Flink no Amazon EMR pode usar para melhorar o tempo de reinício de trabalhos durante as operações de recuperação ou ajuste de escala de tarefas.

Tópicos

- Recuperação local de tarefas
- Ponto de verificação incremental genérico baseado em logs
- Recuperação refinada
- Mecanismo de reinício combinado no agendador adaptável

Recuperação local de tarefas

> **Note**
>
> A recuperação local de tarefas é compatível com o Amazon EMR 6.0.0 e versões superiores.

Com os pontos de verificação do Flink, cada tarefa produz um snapshot do próprio estado que o Flink grava em um armazenamento distribuído, como o Amazon S3. Em casos de recuperação, as tarefas restauram o estado no armazenamento distribuído. O armazenamento distribuído oferece tolerância a falhas e pode redistribuir o estado durante o reajuste de escala por ser acessível a todos os nós.

No entanto, um armazenamento distribuído remoto também tem uma desvantagem: todas as tarefas devem ler seu estado de um local remoto na rede. Isso pode resultar em tempos de recuperação longos para estados grandes durante operações de recuperação ou ajuste de escala de tarefas.

Esse problema de tempo de recuperação longo é resolvido pela recuperação local de tarefas. As tarefas gravam seu estado no ponto de verificação em um armazenamento secundário local para a tarefa, como em um disco local. Elas também armazenam seu estado no armazenamento principal, ou no Amazon S3, no nosso caso. Durante a recuperação, o agendador programa as tarefas no mesmo gerenciador de tarefas em que elas foram executadas anteriormente, para que possam se

recuperar do armazenamento de estado local em vez de ler do armazenamento de estado remoto. Para obter mais informações, consulte Task-Local Recovery na Documentação do Apache Flink.

Nossos testes de benchmark com exemplos de trabalhos mostraram que o tempo de recuperação foi reduzido de minutos para alguns segundos com a recuperação local da tarefa ativada.

Para habilitar a recuperação local de tarefas, defina as configurações a seguir no seu arquivo `flink-conf.yaml`. Especifique o valor do intervalo do ponto de verificação em milissegundos.

```
state.backend.local-recovery: true
state.backend: hasmap or rocksdb
state.checkpoints.dir: s3://storage-location-bucket-path/checkpoint
execution.checkpointing.interval: 15000
```

Ponto de verificação incremental genérico baseado em logs

> **Note**
>
> Os pontos de verificação incrementais genéricos baseados em logs são compatíveis com o Amazon EMR 6.10.0 e versões superiores.

O recurso de ponto de verificação incremental genérico baseado em logs foi adicionado ao Flink 1.16 para melhorar a velocidade dos pontos de verificação. Um intervalo de ponto de verificação mais rápido costuma resultar em uma redução do trabalho de recuperação porque menos eventos precisam ser reprocessados após a recuperação. Para obter mais informações, consulte Improving speed and stability of checkpointing with generic log-based incremental checkpoints no blog do Apache Flink.

Com exemplos de trabalhos, nossos testes de comparação mostraram que o tempo do ponto de verificação foi reduzido de minutos para alguns segundos com o ponto de verificação incremental genérico baseado em logs.

Para habilitar os pontos de verificação incrementais genéricos baseados em logs, defina as configurações a seguir no seu arquivo `flink-conf.yaml`. Especifique o valor do intervalo do ponto de verificação em milissegundos.

```
state.backend.changelog.enabled: true
state.backend.changelog.storage: filesystem
dstl.dfs.base-path: s3://bucket-path/changelog
```

```
    state.backend.local-recovery: true
    state.backend: rocksdb
    state.checkpoints.dir: s3://bucket-path/checkpoint
    execution.checkpointing.interval: 15000
```

Recuperação refinada

> **Note**
>
> O suporte à recuperação refinada do agendador padrão está disponível no Amazon EMR 6.0.0 e versões superiores. O suporte à recuperação refinada no agendador adaptável está disponível no Amazon EMR 6.15.0 e versões superiores.

Quando uma tarefa falha durante a execução, o Flink redefine todo o gráfico de execução e aciona a reexecução completa com base no último ponto de verificação concluído. Isso é mais caro do que apenas reexecutar as tarefas que falharam. A recuperação refinada reinicia somente o componente conectado ao pipeline da tarefa que falhou. No exemplo a seguir, o gráfico do trabalho tem cinco vértices (de `A` a `E`). É feito o pipeline de todas as conexões entre os vértices com distribuição pontual, e o `parallelism.default` do trabalho é definido como `2`.

```
A # B # C # D # E
```

Neste exemplo, há um total de dez tarefas em execução. O primeiro pipeline (`a1toe1`) é executado em um TaskManager (`TM1`), e o segundo pipeline (`a2toe2`) é executado em outro TaskManager (`TM2`).

```
a1 # b1 # c1 # d1 # e1
a2 # b2 # c2 # d2 # e2
```

Há dois componentes conectados por pipeline: `a1 # e1` e `a2 # e2`. Se uma delas `TM1` ou `TM2` falhar, a falha afetará apenas as 5 tarefas no pipeline em que o TaskManager estava correndo. A estratégia de reinicialização só inicia o componente do pipeline afetado.

A recuperação refinada funciona somente com trabalhos perfeitamente paralelos do Flink. Não é compatível com operações de `keyBy()` ou `redistribute()`. Para obter mais informações, consulte FLIP-1: Fine Grained Recovery from Task Failures no projeto Jira Flink Improvement Proposal.

Para habilitar a recuperação refinada, defina as configurações a seguir no seu arquivo `flink-conf.yaml`.

```
jobmanager.execution.failover-strategy: region
restart-strategy: exponential-delay or fixed-delay
```

Mecanismo de reinício combinado no agendador adaptável

> **Note**
>
> O mecanismo de reinício combinado no agendador adaptável é compatível com o Amazon EMR 6.15.0 e versões superiores.

O agendador adaptável pode ajustar o paralelismo do trabalho com base nos slots disponíveis. Ele reduz automaticamente o paralelismo se não houver slots suficientes disponíveis para atender ao paralelismo do trabalho configurado. Se novos slots ficarem disponíveis, o trabalho aumentará a escala verticalmente mais uma vez para o paralelismo do trabalho configurado. Um agendador adaptável evita o tempo de inatividade no trabalho quando não há recursos suficientes disponíveis. Esse é o agendador compatível com o escalador automático do Flink. Recomendamos o agendador adaptável com o Flink no Amazon EMR por esses motivos. No entanto, os agendadores adaptáveis podem fazer várias reinicializações em um curto período; uma reinicialização para cada novo recurso adicionado. Isso pode levar a uma queda na performance do trabalho.

Com o Amazon EMR 6.15.0 e versões superiores, o Flink tem um mecanismo de reinício combinado no agendador adaptável que abre uma janela de reinicialização quando o primeiro recurso é adicionado e aguarda pelo intervalo configurado da janela do padrão de um minuto. Ele executa uma única reinicialização quando há recursos suficientes disponíveis para executar o trabalho com o paralelismo configurado ou quando o intervalo expira.

Com os exemplos de trabalhos, nossos testes de comparação mostraram que esse recurso processa 10% a mais dos registros do que o comportamento padrão ao usar o agendador adaptável e o escalador automático do Flink.

Para habilitar o mecanismo de reinício combinado, defina as configurações a seguir no seu arquivo `flink-conf.yaml`.

```
jobmanager.adaptive-scheduler.combined-restart.enabled: true
jobmanager.adaptive-scheduler.combined-restart.window-interval: 1m
```

Usar trabalhos do Flink pelo Zeppelin no Amazon EMR

Introdução

O Amazon EMR 6.10.0 e versões posteriores oferecem suporte à integração de Apache Zeppelin com o Apache Flink. É possível enviar trabalhos do Flink de forma interativa por meio dos cadernos Zeppelin. Com o intérprete Flink, é possível executar consultas do Flink, definir trabalhos em lote e de transmissão do Flink e visualizar a saída nos cadernos Zeppelin. O intérprete Flink foi desenvolvido com base na API REST do Flink. Isso possibilita acessar e manipular trabalhos do Flink de dentro do ambiente Zeppelin para realizar processamento e análise de dados em tempo real.

Há quatro subintérpretes no intérprete Flink. Eles servem a outras finalidades, mas estão todos na JVM e compartilham os mesmos pontos de entrada pré-configurados do Flink (`ExecutionEnviroment`, `StreamExecutionEnvironment`, `BatchTableEnvironment`, `StreamTableEnvironment`). Os intérpretes são estes:

- `%flink`: cria `ExecutionEnvironment`, `StreamExecutionEnvironment`, `BatchTableEnvironment`, `StreamTableEnvironment` e fornece um ambiente Scala
- `%flink.pyflink`: fornece um ambiente Python
- `%flink.ssql`: fornece um ambiente SQL de transmissão
- `%flink.bsql`: fornece um ambiente SQL em lote

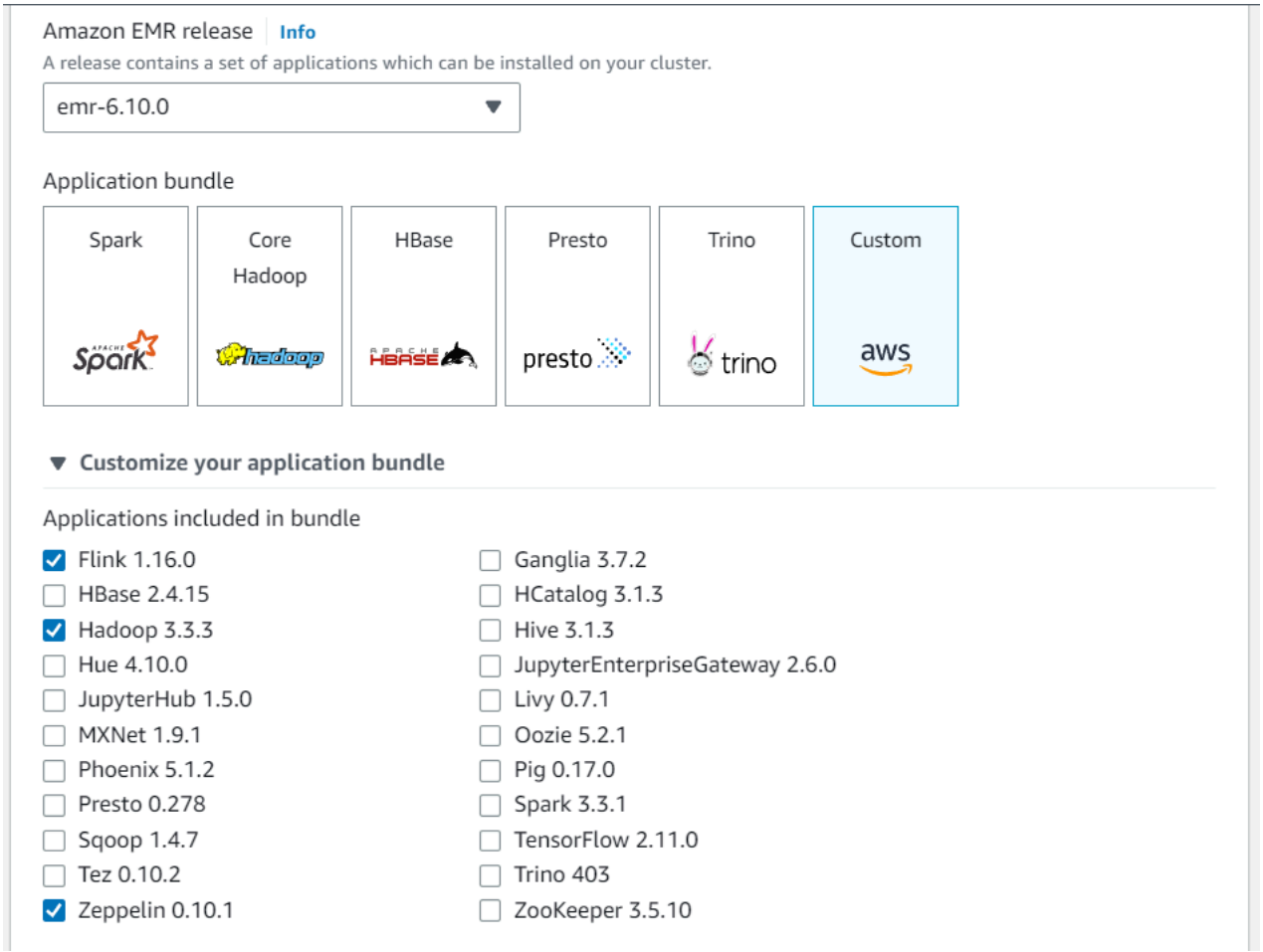
Pré-requisitos

- A integração do Zeppelin com o Flink é compatível com clusters criados com o Amazon EMR 6.10.0 e versões posteriores.
- Para visualizar as interfaces Web hospedadas em clusters do EMR conforme necessário para essas etapas, é necessário configurar um túnel SSH para permitir o acesso de entrada. Para obter mais informações, consulte Configure proxy settings to view websites hosted on the primary node.

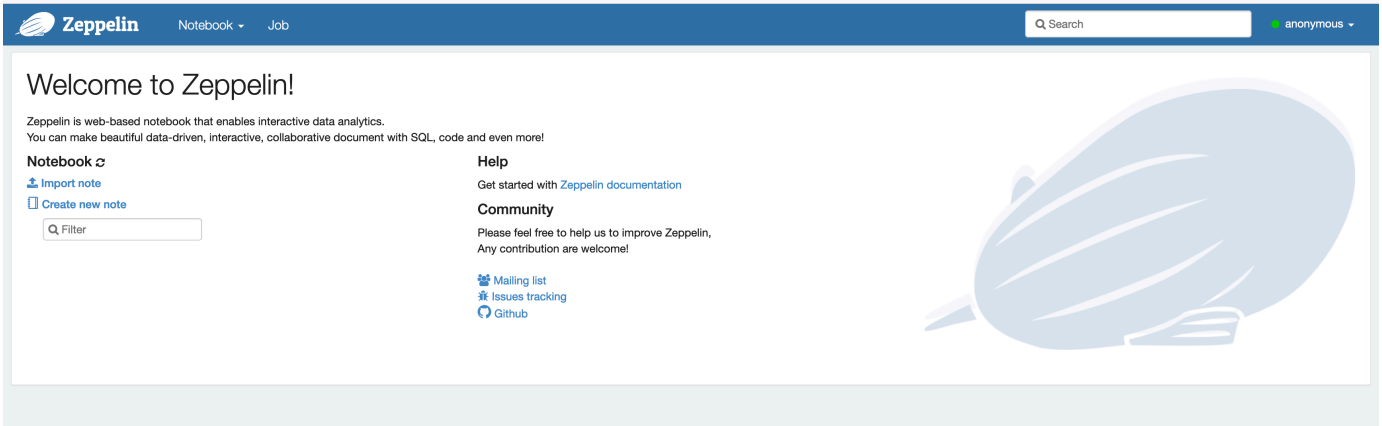
Configurar o Zeppelin-Flink em um cluster do EMR

Realize as etapas a seguir para configurar o Apache Flink no Apache Zeppelin para ser executado em um cluster do EMR:

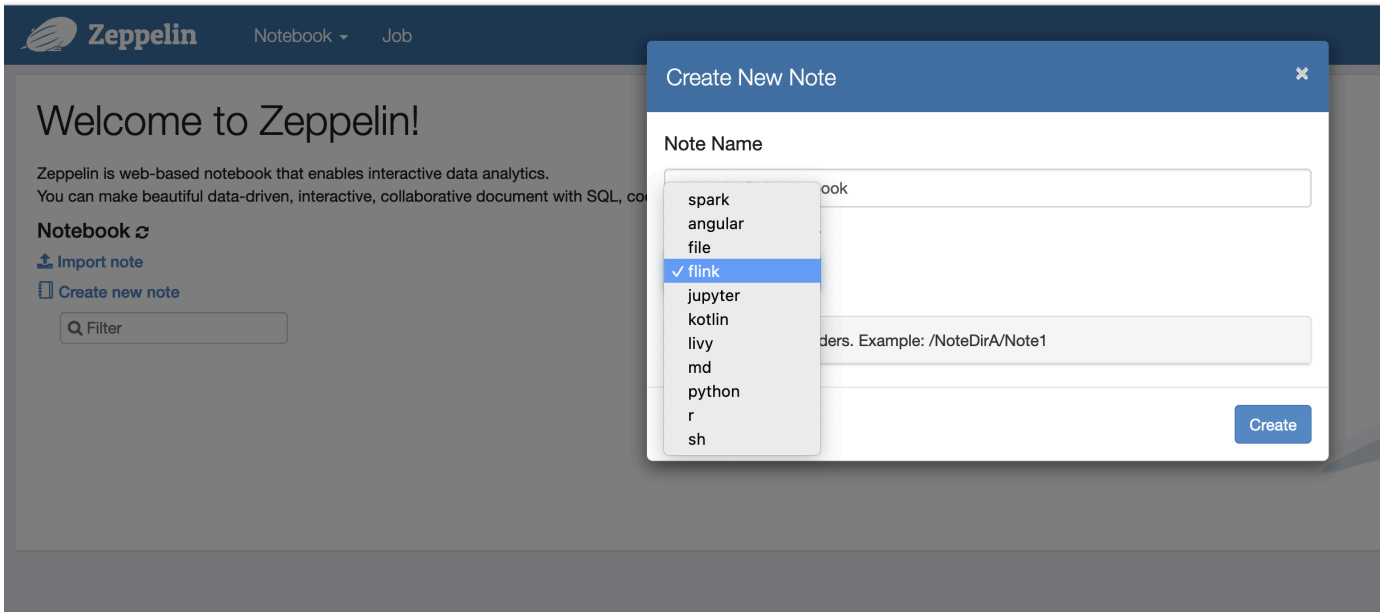
1. Crie um novo cluster pelo console do Amazon EMR. Selecione a versão emr-6.10.0 ou posterior do Amazon EMR. Em seguida, escolha personalizar seu pacote de aplicações com a opção Personalizado. Inclua, no mínimo, Flink, Hadoop e Zeppelin no pacote.



2. Crie o resto do cluster com as configurações de sua preferência.
3. Quando o cluster estiver em execução, selecione o cluster no console para visualizar os detalhes e abra a guia Aplicações. Selecione Zeppelin na seção Interfaces de usuário da aplicação para abrir a interface da Web do Zeppelin. Configure o acesso à interface da Web do Zeppelin com um túnel SSH para o nó primário e uma conexão proxy, conforme descrito em Pré-requisitos.



4. Agora, você pode criar uma nova nota em um caderno do Zeppelin usando o Flink como intérprete padrão.



5. Consulte os exemplos de código a seguir que demonstram como executar trabalhos do Flink em um caderno do Zeppelin.

Executar trabalhos do Flink com o Zeppelin-Flink em um cluster do EMR

- Exemplo 1, Flink Scala

  a) WordCount Exemplo de Batch (SCALA)

```
%flink

val data = benv.fromElements("hello world", "hello flink", "hello hadoop")
```

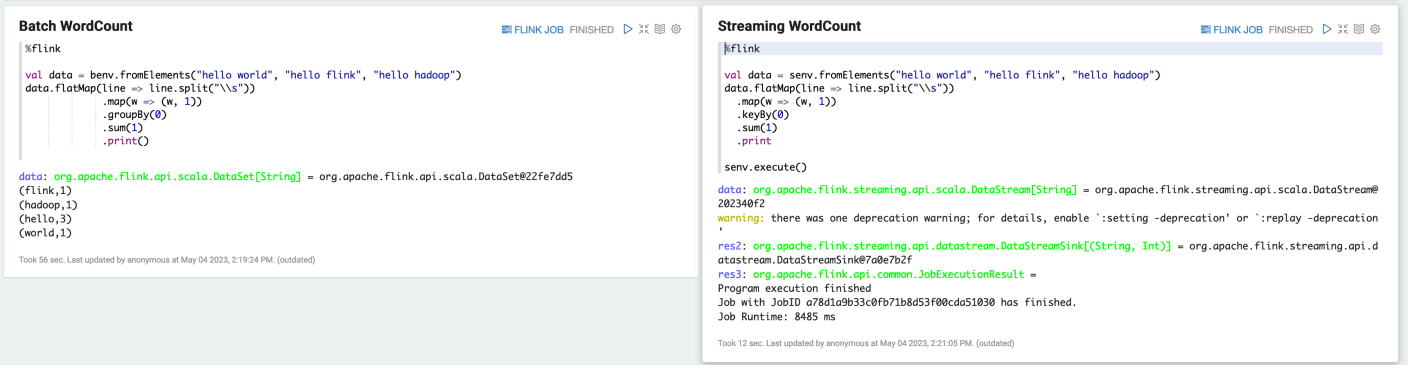
```
data.flatMap(line => line.split("\\s"))
             .map(w => (w, 1))
             .groupBy(0)
             .sum(1)
             .print()
```

b) WordCount Exemplo de streaming (SCALA)

```
%flink

val data = senv.fromElements("hello world", "hello flink", "hello hadoop")
data.flatMap(line => line.split("\\s"))
  .map(w => (w, 1))
  .keyBy(0)
  .sum(1)
  .print

senv.execute()
```



- Exemplo 2, Flink Streaming SQL

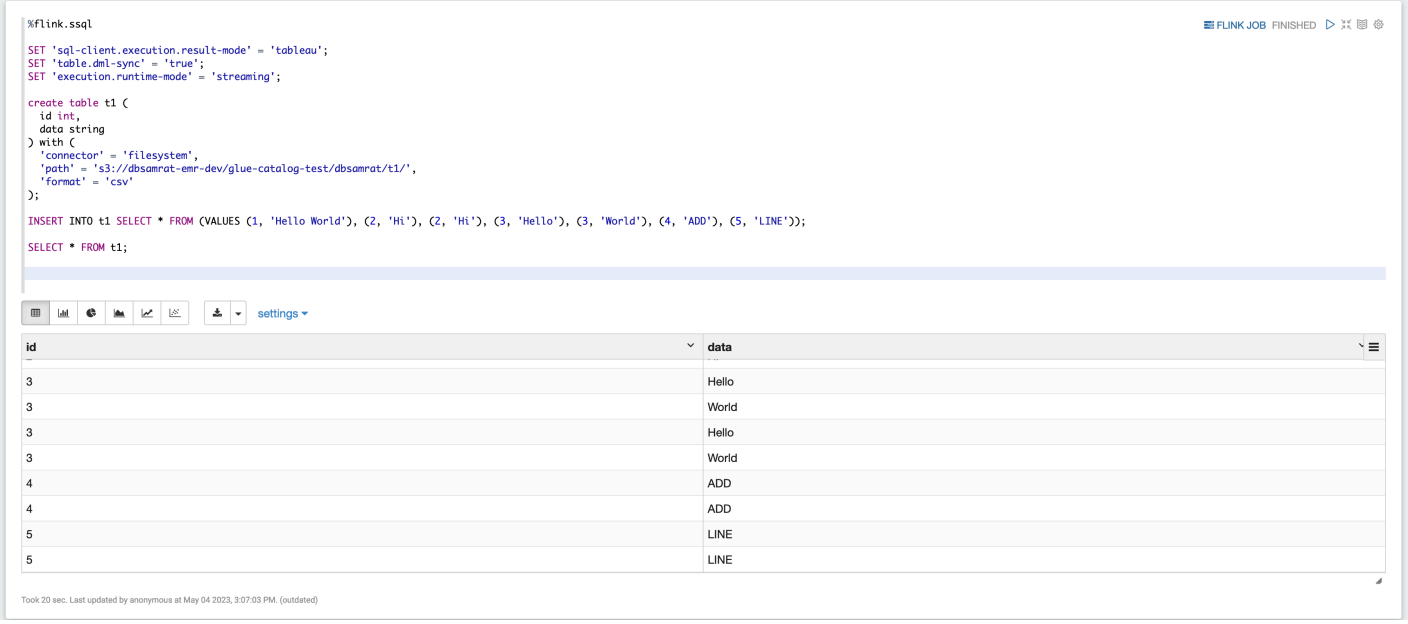
```
%flink.ssql
SET 'sql-client.execution.result-mode' = 'tableau';
SET 'table.dml-sync' = 'true';
SET 'execution.runtime-mode' = 'streaming';

create table dummy_table (
  id int,
  data string
) with (
  'connector' = 'filesystem',
  'path' = 's3://s3-bucket/dummy_table',
  'format' = 'csv'
```

```
);

INSERT INTO dummy_table SELECT * FROM (VALUES (1, 'Hello World'), (2, 'Hi'), (2,
 'Hi'), (3, 'Hello'), (3, 'World'), (4, 'ADD'), (5, 'LINE'));

SELECT * FROM dummy_table;
```



- Exemplo 3, Pyflink. Observe que você deve fazer upload da sua própria amostra de arquivo de texto chamada `word.txt` no bucket do S3.

```
%flink.pyflink

import argparse
import logging
import sys

from pyflink.common import Row
from pyflink.table import (EnvironmentSettings, TableEnvironment, TableDescriptor,
 Schema,
                           DataTypes, FormatDescriptor)
from pyflink.table.expressions import lit, col
from pyflink.table.udf import udtf

def word_count(input_path, output_path):
    t_env = TableEnvironment.create(EnvironmentSettings.in_streaming_mode())
    # write all the data to one file
    t_env.get_config().set("parallelism.default", "1")
```

```
    # define the source
    if input_path is not None:
        t_env.create_temporary_table(
            'source',
            TableDescriptor.for_connector('filesystem')
                           .schema(Schema.new_builder()
                                   .column('word', DataTypes.STRING())
                                   .build())
                           .option('path', input_path)
                           .format('csv')
                           .build())
        tab = t_env.from_path('source')
    else:
        print("Executing word_count example with default input data set.")
        print("Use --input to specify file input.")
        tab = t_env.from_elements(map(lambda i: (i,), word_count_data),
                                  DataTypes.ROW([DataTypes.FIELD('line',
DataTypes.STRING())]))

    # define the sink
    if output_path is not None:
        t_env.create_temporary_table(
            'sink',
            TableDescriptor.for_connector('filesystem')
                           .schema(Schema.new_builder()
                                   .column('word', DataTypes.STRING())
                                   .column('count', DataTypes.BIGINT())
                                   .build())
                           .option('path', output_path)
                           .format(FormatDescriptor.for_format('canal-json')
                                   .build())
                           .build())
    else:
        print("Printing result to stdout. Use --output to specify output path.")
        t_env.create_temporary_table(
            'sink',
            TableDescriptor.for_connector('print')
                           .schema(Schema.new_builder()
                                   .column('word', DataTypes.STRING())
                                   .column('count', DataTypes.BIGINT())
                                   .build())
                           .build())
```

```
    @udtf(result_types=[DataTypes.STRING()])
    def split(line: Row):
        for s in line[0].split():
            yield Row(s)

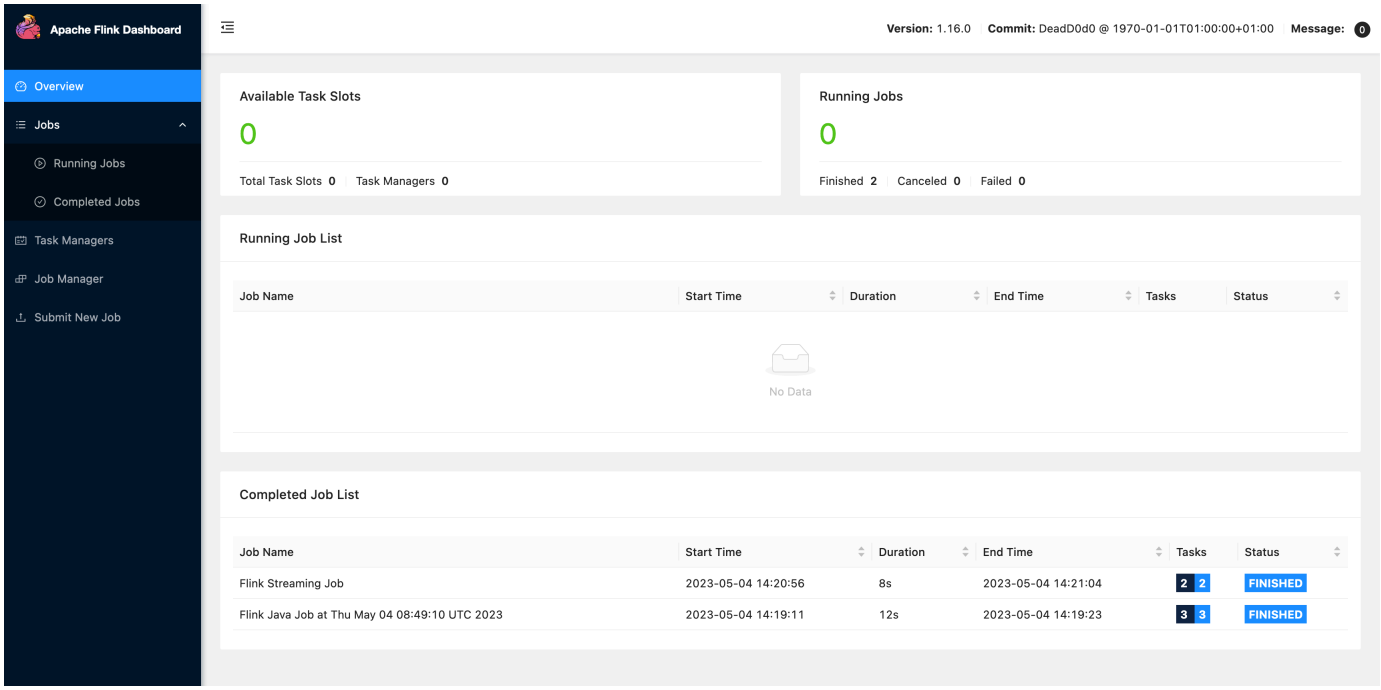
    # compute word count
    tab.flat_map(split).alias('word') \
       .group_by(col('word')) \
       .select(col('word'), lit(1).count) \
       .execute_insert('sink') \
       .wait()


logging.basicConfig(stream=sys.stdout, level=logging.INFO, format="%(message)s")


word_count("s3://s3_bucket/word.txt", "s3://s3_bucket/demo_output.txt")
```

1. Escolha TRABALHO DO FLINK na interface do usuário do Zeppelin para acessar e visualizar a interface do usuário Web do Flink.

2. Escolher TRABALHO DO FLINK encaminhará o console Web do Flink para outra guia do navegador.



Histórico de versões do Flink

A tabela a seguir informa a versão do Flink incluída em cada versão do Amazon EMR, junto com os componentes instalados com a aplicação. Para obter as versões dos componentes em cada versão, consulte a seção Versão de componentes da sua versão em Versões 7.x do Amazon EMR, Versões de lançamento 6.x do Amazon EMR ou Versões de lançamento 5.x do Amazon EMR.

Informações de versão do Flink

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-7.9.0	1.20.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client,

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
		flink-jobmanager-config, hudi, delta
emr-7.8.0	1.20.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta
emr-7.7.0	1.20.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-7.6.0	1.20.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta
emr-7.5.0	1.19.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-7.4.0	1.19.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta
emr-7.3.0	1.18.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-7.2.0	1.18.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta
emr-5.36.2	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-7.1.0	1.18.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-7.0.0	1.18.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors
emr-6.15.0	1.17.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-6.14.0	1.17.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors
emr-6.13.0	1.17.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-6.12.0	1.17.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors
emr-6.11.1	1.16.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-6.11.0	1.16.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors
emr-6.10.1	1.16.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.10.0	1.16.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-6.9.1	1.15.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.9.0	1.15.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.8.1	1.15.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-6.8.0	1.15.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.7.0	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-5.36.1	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.36.0	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-6.6.0	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-5.35.0	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-6.5.0	1.14.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.4.0	1.13.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.3.1	1.12.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-6.3.0	1.12.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-6.2.1	1.11.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-6.2.0	1.11.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-6.1.1	1.11.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-6.1.0	1.11.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.34.0	1.13.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.33.1	1.12.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-5.33.0	1.12.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-5.32.1	1.11.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.32.0	1.11.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-5.31.1	1.11.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-5.31.0	1.11.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.30.2	1.10.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.30.1	1.10.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.30.0	1.10.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.29.0	1.9.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.28.1	1.9.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.28.0	1.9.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.27.1	1.8.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.27.0	1.8.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.26.0	1.8.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.25.0	1.8.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.24.1	1.8.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.24.0	1.8.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.23.1	1.7.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.23.0	1.7.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.22.0	1.7.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.21.2	1.7.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.21.1	1.7.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.21.0	1.7.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.20.1	1.6.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.20.0	1.6.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.19.1	1.6.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.19.0	1.6.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.18.1	1.6.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.18.0	1.6.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.17.2	1.5.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.17.1	1.5.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.17.0	1.5.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.16.1	1.5.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.16.0	1.5.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.15.1	1.4.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.15.0	1.4.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.14.2	1.4.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.14.1	1.4.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.14.0	1.4.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.13.1	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.13.0	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.12.3	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.12.2	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.12.1	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.12.0	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.11.4	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.11.3	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.11.2	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.11.1	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.11.0	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.10.1	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.10.0	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.9.1	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.9.0	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.8.3	1.3.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.8.2	1.3.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.8.1	1.3.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.8.0	1.3.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.7.1	1.3.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.7.0	1.3.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.6.1	1.2.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.6.0	1.2.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.5.4	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.5.3	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.5.2	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.5.1	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.5.0	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.4.1	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.4.0	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.3.2	1.1.4	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.3.1	1.1.4	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.3.0	1.1.4	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.2.3	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.2.2	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.2.1	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.2.0	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.1.1	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

Rótulo de versão do Amazon EMR	Versão do Flink	Componentes instalados com o Flink
emr-5.1.0	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

Notas da versão do Flink por versão

Consulte as seções a seguir para ver as notas de lançamento completas.

Amazon EMR 7.9.0 — Notas de lançamento do Flink

Amazon EMR 7.9.0 - Alterações do Flink

Tipo	Descrição
Novo recurso	A partir do Amazon EMR 7.9.0, o Apache Flink oferece out-of-the-box suporte aos formatos de arquivo Avro, Parquet e ORC. Você pode usar esses formatos diretamente com qualquer API do Flink (DataStream, tabela ou SQL) sem precisar de nenhuma configuração adicional.
Novo recurso	A partir da versão 7.9.0 do Amazon EMR, você pode habilitar os catálogos de dados Hive Metastore ou AWS Glue com mais facilidade usando as definições de configuração. Adicione uma `hive.enabled: true` ou mais `glue.enabled: true` na `flink-conf` classificação durante a criação do cluster para configurar automaticamente o respectivo

Tipo	Descrição
	catálogo de dados. Essa abordagem simplific ada elimina as etapas de configuração manual que eram necessárias anteriormente.

Amazon EMR 7.8.0 — Notas de lançamento do Flink

Configuração — O EMR Flink funciona imediatamente com o S3A em todas as regiões/partições. AWS

Amazon EMR 7.7.0 — Notas de lançamento do Flink

- O shell SQL do Flink pode ser invocado facilmente com o comando `flink-sql-client` que está vinculado a `/usr/lib/flink/bin/sql-client.sh`

Amazon EMR 7.6.0 — Notas de lançamento do Flink

Amazon EMR 7.6.0 - Características do Flink

- Nenhuma alteração no lançamento.

Amazon EMR 7.5.0 — Notas de lançamento do Flink

Tipo	Descrição
Recurso	Foi adicionado suporte para executar trabalhos do Flink com um jar remoto.
Melhoria	Torne o encadeamento de exclusão e inclusão de vértices seguro.

Amazon EMR 7.5.0 - Características do Flink

- A partir do Amazon EMR 7.5.0, você pode especificar uma localização do Amazon S3 como o caminho do JAR ao usar os comandos da CLI e do `run` `run-application` Apache Flink. Quando você fornece um caminho do S3, o EMR baixa automaticamente o arquivo JAR do

Amazon S3 para o armazenamento do EBS do cluster. Sempre que você especifica o mesmo arquivo JAR, o EMR baixa a versão mais recente do Amazon S3 em vez de reutilizar o arquivo JAR existente no cluster.

- A partir do Amazon EMR 7.5.0, os clientes podem passar o caminho remoto (um local do S3) como o caminho JAR com os comandos da CLI do `run` `run-application` Flink. O JAR é então automaticamente retirado do armazenamento S3 para o armazenamento EBS do cluster. Se o mesmo JAR for fornecido novamente, ele baixará o mais recente do S3 e não reutilizará o JAR existente no cluster.

Amazon EMR 7.4.0 — Notas de lançamento do Flink

Tipo	Descrição
Upgrade	A versão Flink foi atualizada para 1.19.1.

Amazon EMR 7.3.0: notas de versão do Flink

- Por padrão, clusters com criptografia em trânsito habilitada por meio da configuração de segurança usarão o TLS 1.3 para comunicação interna entre os processos do Flink, o endpoint REST do gerenciador de trabalhos e o servidor de histórico de trabalho do Flink.

Amazon EMR 7.2.0: notas de versão do Flink

Tipo	Descrição
Melhoria	Suporte à adição de rótulos personalizados ao serviço Kubernetes por trabalho do Flink por meio da configuração `kubernete s.service.labels` .

Ganglia

> **Note**
>
> A última versão do Amazon EMR a incluir o Ganglia foi o Amazon EMR 6.15.0. Para monitorar seu cluster, versões superiores à 6.15.0 incluem o agente Amazon CloudWatch .

O projeto de código-fonte aberto Ganglia é um sistema distribuído e escalável, projetado para monitorar clusters e grades e, ao mesmo tempo, minimizar o impacto sobre o seu desempenho. Quando você habilita o Ganglia no seu cluster, pode gerar relatórios e visualizar o desempenho do cluster como um todo, bem como inspecionar o desempenho de instâncias de nós individuais. O Ganglia também é configurado para analisar e visualizar as métricas do Hadoop e do Spark. Para obter mais informações sobre o projeto de código-fonte aberto Ganglia, acesse http://ganglia.info/.

Ao visualizar a interface do usuário da Web do Ganglia em um navegador, você tem uma visão geral da performance do cluster, com grafos detalhando a carga, o uso da memória, a utilização da CPU e o tráfego de rede do cluster. Abaixo das estatísticas do cluster estão gráficos para cada servidor do cluster.

A tabela a seguir lista a versão do Ganglia incluída na versão mais recente da série 6.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Ganglia.

Para obter a versão dos componentes instalados com o Ganglia nessa versão, consulte Release 6.15.0 Component Versions.

Informações de versão do Ganglia para o emr-6.15.0

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-6.15.0	Ganglia 3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodema

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
		nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server, webserver

A tabela a seguir lista a versão do Ganglia incluída na versão mais recente da série 5.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Ganglia.

Para obter a versão dos componentes instalados com o Ganglia nesse lançamento, consulte Release 5.36.2 Component Versions.

Informações de versão do Ganglia para o emr-5.36.2

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.36.2	Ganglia 3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Tópicos

- Criar um cluster com o Ganglia
- Visualizar métricas do Ganglia
- Métricas do Hadoop e do Spark no Ganglia
- Histórico de versões do Ganglia

Criar um cluster com o Ganglia

> **Note**
>
> A última versão do Amazon EMR a incluir o Ganglia foi o Amazon EMR 6.15.0. Para monitorar seu cluster, versões superiores à 6.15.0 incluem o agente Amazon CloudWatch .

Para criar um cluster com o Ganglia usando o console

1. Navegue até o novo console do Amazon EMR e selecione Alternar para o console antigo na navegação lateral. Para obter mais informações sobre o que esperar ao alternar para o console antigo, consulte Usar o console antigo.
2. Selecione Criar cluster.
3. Em Software configuration (Configuração do software), escolha All Applications (Todos os aplicativos), Core Hadoop ou Spark.
4. Prossiga com a criação do cluster usando as configurações apropriadas.

Para adicionar o Ganglia a um cluster usando o AWS CLI

No AWS CLI, você pode adicionar o Ganglia a um cluster usando `create-cluster` com o `--applications` parâmetro. Se você especificar apenas o Ganglia usando o parâmetro `--applications`, o Ganglia será o único aplicativo instalado.

- Digite o comando a seguir para adicionar o Ganglia ao criar um cluster e *myKey* substituí-lo pelo nome do seu EC2 key pair.

  > **Note**
  >
  > Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr create-cluster --name "Spark cluster with Ganglia" --release-label
 emr-6.15.0; \
--applications Name=Spark Name=Ganglia \
```

```
--ec2-attributes KeyName=myKey --instance-type m5.xlarge \
--instance-count 3 --use-default-roles
```

Quando você especifica a contagem de instâncias sem usar o parâmetro `--instance-groups`, um único nó principal é executado, e as instâncias restantes são executadas como nós core. Todos os nós usam o tipo de instância especificado no comando.

> **Note**
>
> Se você ainda não criou a função de serviço e o perfil de EC2 instância padrão do EMR, digite `aws emr create-default-roles` para criá-los antes de digitar o subcomando. `create-cluster`

Para obter mais informações sobre o uso dos comandos do Amazon EMR no AWS CLI, consulte. https://docs.aws.amazon.com/cli/latest/reference/emr

Visualizar métricas do Ganglia

> **Note**
>
> A última versão do Amazon EMR a incluir o Ganglia foi o Amazon EMR 6.15.0. Para monitorar seu cluster, versões superiores à 6.15.0 incluem o agente Amazon CloudWatch .

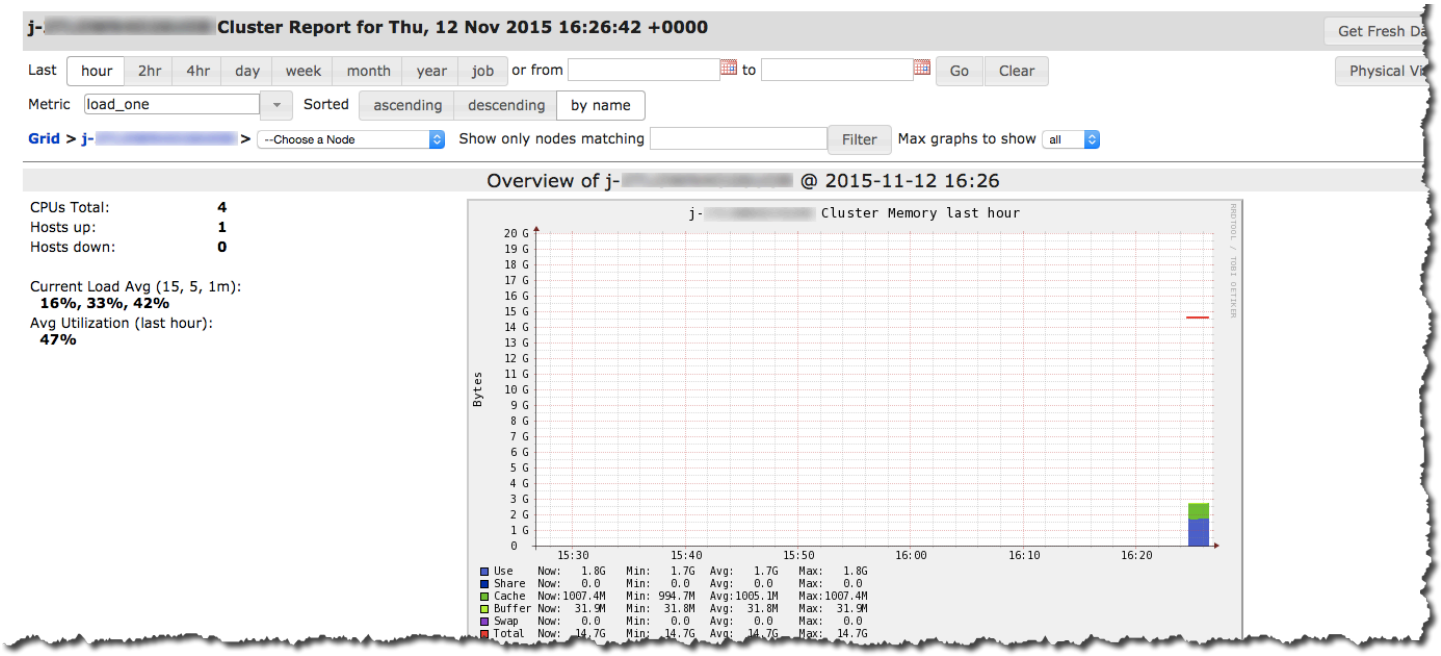
O Ganglia fornece uma interface de usuário na web que você pode usar para visualizar as métricas que ele coleta. Quando você executa o Ganglia no Amazon EMR, a interface Web é executada no nó principal e pode ser visualizada por meio do encaminhamento de porta, também conhecido como criação de um túnel de SSH. Para obter mais informações sobre a visualização de interfaces Web no Amazon EMR, consulte View web interfaces hosted on EMR clusters no Guia de gerenciamento do Amazon EMR.

Para visualizar a interface da web do Ganglia

1. Use o SSH para formar um túnel até o nó principal e criar uma conexão segura. Para obter informações sobre como criar um túnel SSH até o nó principal, consulte a Option 2, part 1: Set

up an SSH tunnel to the master node using dynamic port forwarding no Guia de gerenciamento do Amazon EMR.

2. Instale um navegador da Web com uma ferramenta de proxy, como o FoxyProxy plug-in do Firefox, para criar um proxy SOCKS para domínios do tipo *ec2*.amazonaws.com*. Para obter mais informações, consulte Option 2, part 2: Configure proxy settings to view websites hosted on the master node no Guia de gerenciamento do Amazon EMR.
3. Com o proxy definido e a conexão SSH aberta, você pode visualizar a interface do usuário do Ganglia abrindo uma janela do navegador com http://*master-public-dns-name*/ganglia/, onde *master-public-dns-name* está o endereço DNS público do servidor mestre no cluster do EMR.



Métricas do Hadoop e do Spark no Ganglia

> **Note**
>
> A última versão do Amazon EMR a incluir o Ganglia foi o Amazon EMR 6.15.0. Para monitorar seu cluster, versões superiores à 6.15.0 incluem o agente Amazon CloudWatch .

O Ganglia informa as métricas do Hadoop para cada instância. Os vários tipos de métricas são prefixados por categoria: sistema de arquivos distribuído (dfs.*), máquina virtual Java (jvm.*), (mapred.*) e chamadas de procedimento remoto MapReduce (rpc.*).

As métricas do Ganglia com base no YARN, como o Spark e o Hadoop, não estão disponíveis para as versões 4.4.0 e 4.5.0 do EMR. Use uma versão mais recente para usar essas métricas.

As métricas do Ganglia para o Spark geralmente têm prefixos para o ID do aplicativo YARN e o Spark. DAGScheduler Portanto, os prefixos sigam este formato:

- DAGScheduler.*
- application_xxxxxxxxxx_xxxx.driver.*
- application_xxxxxxxxxx_xxxx.executor.*

Histórico de versões do Ganglia

> **Note**
>
> A última versão do Amazon EMR a incluir o Ganglia foi o Amazon EMR 6.15.0. Para monitorar seu cluster, versões superiores à 6.15.0 incluem o agente Amazon CloudWatch .

A tabela a seguir informa a versão do Ganglia incluída em cada versão do Amazon EMR, junto com os componentes instalados com a aplicação. Para obter as versões dos componentes em cada versão, consulte a seção Versão de componentes da sua versão em Versões 7.x do Amazon EMR, Versões de lançamento 6.x do Amazon EMR ou Versões de lançamento 5.x do Amazon EMR.

Informações sobre a versão do Ganglia

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.36.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
		server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.15.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.14.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-6.13.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.12.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.11.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-6.11.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.10.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.10.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-6.9.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.9.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.8.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-6.8.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.7.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.36.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.36.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.6.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.35.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-6.5.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.4.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.3.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-6.3.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.2.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.2.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-6.1.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.1.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.0.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-6.0.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.34.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.33.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.33.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.32.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.32.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.31.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.31.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.30.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.30.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.30.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.29.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.28.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.28.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.27.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.27.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.26.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.25.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.24.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.24.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.23.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.23.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.22.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.21.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.21.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.21.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.20.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.20.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.19.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.19.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.18.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.18.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.17.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.17.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.17.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.16.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.16.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.15.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.15.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.14.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.14.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.14.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.13.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.13.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.12.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.12.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.12.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.12.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.11.4	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.11.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.11.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.11.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.11.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.10.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.10.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.9.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.9.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.8.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.8.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.8.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.8.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.7.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.7.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.6.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.6.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.5.4	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.5.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.5.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.5.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.5.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.4.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.4.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.3.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.3.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.3.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.2.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.2.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.2.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.2.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.1.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.1.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.0.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.0.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.0.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-5.0.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.9.6	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.9.5	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-4.9.4	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.9.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.9.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-4.9.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.8.5	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.8.4	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-4.8.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.8.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.8.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-4.8.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.7.4	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.7.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-4.7.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.7.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.7.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-4.6.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.6.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.5.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Rótulo de versão do Amazon EMR	Versão do Ganglia	Componentes instalados com o Ganglia
emr-4.4.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.3.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.2.0	3.6.0	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Apache Hadoop

O Apache Hadoop é uma estrutura de software Java de código aberto compatível com processamento de dados em massa por todo um cluster de instâncias. Ele também pode ser executado em uma única instância ou em milhares de instâncias. O Hadoop usa muitos modelos de processamento, como o Tez MapReduce e o Tez, para distribuir o processamento em várias instâncias e também usa um sistema de arquivos distribuído chamado HDFS para armazenar dados em várias instâncias. O Hadoop monitora a integridade das instâncias no cluster e pode fazer a recuperação de falha em um ou mais nós. Dessa maneira, o Hadoop proporciona o aumento do processamento e da capacidade de armazenamento, bem como a alta disponibilidade. Para obter mais informações, consulte a documentação do Hadoop.

A tabela a seguir lista a versão do Hadoop incluída na versão mais recente da série 7.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Hadoop.

Para a versão dos componentes instalados com o Hadoop nesta versão, consulte Versões de componentes da versão 7.9.0.

Informações sobre a versão do Hadoop para o emr-7.9.0

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-7.9.0	Hadoop 3.4.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

A tabela a seguir lista a versão do Hadoop incluída na versão mais recente da série 6.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Hadoop.

Para obter a versão dos componentes instalados com o Hadoop nessa versão, consulte Release 6.15.0 Component Versions.

Informações de versão do Hadoop para o emr-6.15.0

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-6.15.0	Hadoop 3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

A tabela a seguir lista a versão do Hadoop incluída na versão mais recente da série 5.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Hadoop.

Para obter a versão dos componentes instalados com o Hadoop nesse lançamento, consulte Release 5.36.2 Component Versions.

Informações de versão do Hadoop para o emr-5.36.2

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.36.2	Hadoop 2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodema

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
		nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server

Desde a versão 5.18.0 do Amazon EMR, você pode usar o repositório de artefatos do Amazon EMR para criar o código de trabalho em comparação com as versões exatas de bibliotecas e dependências disponíveis com versões específicas do Amazon EMR. Para obter mais informações, consulte Verificar dependências usando o repositório de artefatos do Amazon EMR.

Tópicos

- Configurar o Hadoop
- Criptografia transparente no HDFS com o Amazon EMR
- Criar ou executar uma aplicação Hadoop
- Ler objetos restaurados
- Ative o reconhecimento de acesso não uniforme à memória para contêineres do YARN
- Embalagem de contêiner YARN
- Histórico de versões do Hadoop

Configurar o Hadoop

As seções a seguir oferecem as definições de configuração padrão para os daemons do Hadoop, a tarefas e o HDFS.

Tópicos

- Configuração da tarefa
- Definições de configuração de daemon do Hadoop
- Configuração do HDFS

Configuração da tarefa

Você pode definir variáveis de configuração para ajustar o desempenho de seus MapReduce trabalhos. Esta seção fornece os valores padrão para configurações importantes. Os valores padrão variam com base no tipo de EC2 instância do nó usado no cluster. HBase está disponível ao usar a versão 4.6.0 e posterior do Amazon EMR. Diferentes padrões são usados quando HBase é instalado. Esses valores são fornecidos juntamente com os padrões iniciais.

O Hadoop 2 usa dois parâmetros, `mapreduce.map.java.opts` e`mapreduce.reduce.java.opts`, para configurar a memória para mapear e reduzirJVMs , respectivamente. Eles substituem a opção de configuração única `mapreduce.map.java.opts` nas versões anteriores do Hadoop.

Da mesma forma, `mapred.job.jvm.num.tasks` substitui `mapred.job.reuse.jvm.num.tasks` no Hadoop 2.7.2 e versões posteriores. O Amazon EMR define esse valor como 20, independentemente do tipo de EC2 instância. Você pode substituir essa configuração usando a classificação de configuração `mapred-site`. Definir o valor `-1` indica que uma JVM pode ser reutilizada para um número infinito de tarefas em um único trabalho, e o valor 1 indica que uma nova JVM é gerada para cada tarefa.

Por exemplo, para definir o valor de `mapred.job.jvm.num.tasks` como -1, você pode criar um arquivo com o seguinte conteúdo:

```
[
 {
  "Classification": "mapred-site",
  "Properties": {
   "mapred.job.jvm.num.tasks": "-1"
  }
 }
]
```

Ao usar o `create-cluster` comando ou `modify-instance-groups` comando do AWS CLI, você pode então referenciar o arquivo de configuração JSON. No exemplo a seguir, o arquivo de configuração é salvo como `myConfig.json` e armazenado no Amazon S3.

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge \
--instance-count 3 --applications Name=Hadoop --configurations https://
s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myConfig.json \
--use-default-roles
```

Você pode alterar valores padrão listados abaixo usando a classificação de configuração `mapred-site` da mesma forma e definir vários valores e várias classificações de configuração usando um único arquivo JSON. Para obter mais informações, consulte Configurar aplicações.

Com as versões 5.21.0 e posteriores do Amazon EMR, você pode substituir as configurações de cluster e especificar classificações de configuração adicionais para cada grupo de instâncias em um cluster em execução. Você faz isso usando o console do Amazon EMR, o AWS Command Line Interface (AWS CLI) ou o AWS SDK. Para obter mais informações, consulte Supplying a Configuration for an Instance Group in a Running Cluster.

Valores padrão para definições de configuração de tarefa

> **Note**
>
> Os valores especificados para cada instância estão de acordo com a versão mais recente do EMR. As versões anteriores podem ter configurações padrão diferentes.

Tipos de instância

- Instâncias c1
- Instâncias c3
- Instâncias c4
- Instâncias c5

- Instâncias c5a
- Instâncias c5ad
- Instâncias c5d
- Instâncias c5n
- Instâncias c6a
- Instâncias c6g
- Instâncias c6gd
- Instâncias c6gn
- Instâncias c6i
- Instâncias c6id
- Instâncias c6in
- Instâncias c7a
- Instâncias c7g
- Instâncias c7gd
- Instâncias c7gn
- Instâncias C7i
- instâncias c7i-flex
- instâncias c8g
- instâncias c8gd
- Instâncias d2
- Instâncias d3
- Instâncias d3en
- instâncias f2
- Instâncias g3
- Instâncias g3s
- Instâncias do g4dn
- Instâncias g5
- instâncias g6

- instâncias g6e
- instâncias gr6
- Instâncias h1
- Instâncias i2
- Instâncias i3
- Instâncias i3en
- Instâncias i4g
- Instâncias i4i
- instâncias i7i
- instâncias i7ie
- instâncias i8g
- Instâncias im4gn
- Instâncias is4gen
- Instâncias m1
- Instâncias m2
- Instâncias m3
- Instâncias m4
- Instâncias m5
- Instâncias m5a
- Instâncias m5ad
- Instâncias m5d
- Instâncias m5dn
- Instâncias m5n
- instâncias m5zn
- Instâncias m6a
- Instâncias m6g
- Instâncias m6gd
- Instâncias m6i

- Instâncias m6id
- Instâncias m6idn
- Instâncias m6in
- Instâncias M7a
- Instâncias m7g
- Instâncias m7gd
- Instâncias M7i
- Instâncias M7i-flex
- instâncias m8g
- instâncias m8gd
- Instâncias p2
- Instâncias p3
- instâncias p4d
- Instâncias p5
- Instâncias r3
- Instâncias r4
- Instâncias r5
- Instâncias r5a
- Instâncias r5ad
- Instâncias r5b
- Instâncias r5d
- Instâncias r5dn
- Instâncias r5n
- Instâncias r6a
- Instâncias r6g
- Instâncias r6gd
- Instâncias r6i
- Instâncias r6iD
- instâncias r6idn

- Instâncias r6in
- Instâncias R7a
- Instâncias r7g
- Instâncias r7gd
- Instâncias r7i
- Instâncias R7iz
- instâncias r8g
- instâncias r8gd
- Instâncias x1
- Instâncias x1e
- Instâncias x2gd
- instâncias x2idn
- instâncias x2iedn
- instâncias x8g
- Instâncias z1d

Instâncias c1

c1.medium

Opção de configuração	Valor padrão	Com HBase instalado
`mapreduce.map.java.opts`	-Xmx288m	-Xmx288m
`mapreduce.java.opts`	-Xmx288m	-Xmx288m
`mapreduce.map.memory.mb`	512	512
`mapreduce.reduce.memory.mb`	512	512
`yarn.app.mapreduce.am.resource.mb`	512	512
`yarn.scheduler.minimum-allocation-mb`	1	1

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.maximum-allocation-mb	512	512
yarn.nodemanager.resource.memory-mb	1024	512

c1.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx864m	-Xmx864m
mapreduce.java.opts	-Xmx1536m	-Xmx1536m
mapreduce.map.memory.mb	1024	1024
mapreduce.reduce.memory.mb	2048	2048
yarn.app.mapreduce.am.resource.mb	2048	2048
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	2048	2560
yarn.nodemanager.resource.memory-mb	5120	2560

Instâncias c3

c3.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1.408	1.408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c3.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1152m	-Xmx1152m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1440	1440
mapreduce.reduce.memory.mb	2.880	2.880
yarn.app.mapreduce.am.resource.mb	2.880	2.880
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11520	5760

Opção de configuração	Valor padrão	Com HBase instalado
yarn.nodemanager.resource.memory-mb	11520	5760

c3.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1152m	-Xmx1152m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1440	1440
mapreduce.reduce.memory.mb	2.880	2.880
yarn.app.mapreduce.am.resource.mb	2.880	2.880
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23040	11520
yarn.nodemanager.resource.memory-mb	23040	11520

c3.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1331m	-Xmx1331m
mapreduce.java.opts	-Xmx2662m	-Xmx2662m
mapreduce.map.memory.mb	1664	1664

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.reduce.memory.mb	3328	3328
yarn.app.mapreduce.am.resource.mb	3328	3328
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	53248	26624
yarn.nodemanager.resource.memory-mb	53248	26624

Instâncias c4

c4.large

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx717m	-Xmx717m
mapreduce.java.opts	-Xmx1434m	-Xmx1434m
mapreduce.map.memory.mb	896	896
mapreduce.reduce.memory.mb	1792	1792
yarn.app.mapreduce.am.resource.mb	1792	1792
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1792	896

Opção de configuração	Valor padrão	Com HBase instalado
yarn.nodemanager.resource.memory-mb	1792	896

c4.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1.408	1.408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c4.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1152m	-Xmx1152m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1440	1440

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.reduce.memory.mb	2.880	2.880
yarn.app.mapreduce.am.resource.mb	2.880	2.880
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11520	5760
yarn.nodemanager.resource.memory-mb	11520	5760

c4.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1152m	-Xmx1152m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1440	1440
mapreduce.reduce.memory.mb	2.880	2.880
yarn.app.mapreduce.am.resource.mb	2.880	2.880
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23040	11520
yarn.nodemanager.resource.memory-mb	23040	11520

c4.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1183m	-Xmx1183m
mapreduce.java.opts	-Xmx2366m	-Xmx2366m
mapreduce.map.memory.mb	1479	1479
mapreduce.reduce.memory.mb	2958	2958
yarn.app.mapreduce.am.resource.mb	2958	2958
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	53248	26624
yarn.nodemanager.resource.memory-mb	53248	26624

Instâncias c5

c5.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	6144	3072
yarn.nodemanager.resource.memory-mb	6144	3072

c5.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

c5.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

c5.9xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1456m	-Xmx1456m
mapreduce.java.opts	-Xmx2912m	-Xmx2912m
mapreduce.map.memory.mb	1820	1820
mapreduce.reduce.memory.mb	3640	3640
yarn.app.mapreduce.am.resource.mb	3640	3640

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	65536	32768
yarn.nodemanager.resource.memory-mb	65536	32768

c5.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1502m	-Xmx1502m
mapreduce.java.opts	-Xmx3004m	-Xmx3004m
mapreduce.map.memory.mb	1877	1877
mapreduce.reduce.memory.mb	3754	3754
yarn.app.mapreduce.am.resource.mb	3754	3754
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	90112	30048
yarn.nodemanager.resource.memory-mb	90112	30048

c5.18xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1547m	-Xmx1547m
mapreduce.java.opts	-Xmx3094m	-Xmx3094m
mapreduce.map.memory.mb	1934	1934
mapreduce.reduce.memory.mb	3868	3868
yarn.app.mapreduce.am.resource.mb	3868	3868
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	139264	30960
yarn.nodemanager.resource.memory-mb	139264	30960

c5.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1570m	-Xmx1570m
mapreduce.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.map.memory.mb	1963	1963
mapreduce.reduce.memory.mb	3926	3926
yarn.app.mapreduce.am.resource.mb	3926	3926

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31376
yarn.nodemanager.resource.memory-mb	188416	31376

Instâncias c5a

c5a.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1.408	1.408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c5a.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c5a.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c5a.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c5a.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1502m	-Xmx1502m
mapreduce.java.opts	-Xmx3004m	-Xmx3004m
mapreduce.map.memory.mb	1877	1877
mapreduce.reduce.memory.mb	3754	3754
yarn.app.mapreduce.am.resource.mb	3754	3754
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	90112	30048
yarn.nodemanager.resource.memory-mb	90112	30048

c5a.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c5a.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

Instâncias c5ad

c5ad.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1.408	1.408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c5ad.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c5ad.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c5ad.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c5ad.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c5ad.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c5ad.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

Instâncias c5d

c5d.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	6144	3072
yarn.nodemanager.resource.memory-mb	6144	3072

c5d.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

c5d.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

c5d.9xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1456m	-Xmx1456m
mapreduce.java.opts	-Xmx2912m	-Xmx2912m
mapreduce.map.memory.mb	1820	1820
mapreduce.reduce.memory.mb	3640	3640
yarn.app.mapreduce.am.resource.mb	3640	3640

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	65536	32768
yarn.nodemanager.resource.memory-mb	65536	32768

c5d.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1502m	-Xmx1502m
mapreduce.java.opts	-Xmx3004m	-Xmx3004m
mapreduce.map.memory.mb	1877	1877
mapreduce.reduce.memory.mb	3754	3754
yarn.app.mapreduce.am.resource.mb	3754	3754
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	90112	30048
yarn.nodemanager.resource.memory-mb	90112	30048

c5d.18xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1547m	-Xmx1547m
mapreduce.java.opts	-Xmx3094m	-Xmx3094m
mapreduce.map.memory.mb	1934	1934
mapreduce.reduce.memory.mb	3868	3868
yarn.app.mapreduce.am.resource.mb	3868	3868
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	139264	30960
yarn.nodemanager.resource.memory-mb	139264	30960

c5d.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1570m	-Xmx1570m
mapreduce.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.map.memory.mb	1963	1963
mapreduce.reduce.memory.mb	3926	3926
yarn.app.mapreduce.am.resource.mb	3926	3926

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31376
yarn.nodemanager.resource.memory-mb	188416	31376

Instâncias c5n

c5n.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1613m	-Xmx1613m
mapreduce.java.opts	-Xmx3226m	-Xmx3226m
mapreduce.map.memory.mb	2016	2016
mapreduce.reduce.memory.mb	4032	4032
yarn.app.mapreduce.am.resource.mb	4032	4032
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	8064	4032
yarn.nodemanager.resource.memory-mb	8064	4032

c5n.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1613m	-Xmx1613m
mapreduce.java.opts	-Xmx3226m	-Xmx3226m
mapreduce.map.memory.mb	2016	2016
mapreduce.reduce.memory.mb	4032	4032
yarn.app.mapreduce.am.resource.mb	4032	4032
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	16128	8064
yarn.nodemanager.resource.memory-mb	16128	8064

c5n.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1741m	-Xmx1741m
mapreduce.java.opts	-Xmx3482m	-Xmx3482m
mapreduce.map.memory.mb	2176	2176
mapreduce.reduce.memory.mb	4352	4352
yarn.app.mapreduce.am.resource.mb	4352	4352

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	34816	17408
yarn.nodemanager.resource.memory-mb	34816	17408

c5n.9xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2002m	-Xmx2002m
mapreduce.java.opts	-Xmx4004m	-Xmx4004m
mapreduce.map.memory.mb	2503	2503
mapreduce.reduce.memory.mb	5006	5006
yarn.app.mapreduce.am.resource.mb	5006	5006
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	90112	30040
yarn.nodemanager.resource.memory-mb	90112	30040

c5n.18xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2094m	-Xmx2094m
mapreduce.java.opts	-Xmx4188m	-Xmx4188m
mapreduce.map.memory.mb	2617	2617
mapreduce.reduce.memory.mb	5234	5234
yarn.app.mapreduce.am.resource.mb	5234	5234
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31396
yarn.nodemanager.resource.memory-mb	188416	31396

Instâncias c6a

c6a.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1.408	1.408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6a.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c6a.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6a.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c6a.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c6a.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c6a.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c6a.32xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1510m	-Xmx1510m
mapreduce.java.opts	-Xmx3020m	-Xmx3020m
mapreduce.map.memory.mb	1888	1888
mapreduce.reduce.memory.mb	3776	3776
yarn.app.mapreduce.am.resource.mb	3776	3776
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

c6a.48xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1527m	-Xmx1527m
mapreduce.java.opts	-Xmx3054m	-Xmx3054m
mapreduce.map.memory.mb	1909	1909
mapreduce.reduce.memory.mb	3818	3818
yarn.app.mapreduce.am.resource.mb	3818	3818
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30608
yarn.nodemanager.resource.memory-mb	366592	30608

Instâncias c6g

c6g.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1.408	1.408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6g.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c6g.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6g.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c6g.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c6g.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

Instâncias c6gd

c6gd.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1.408	1.408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6gd.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c6gd.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6gd.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c6gd.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c6gd.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

Instâncias c6gn

c6gn.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1.408	1.408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6gn.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c6gn.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6gn.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c6gn.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c6gn.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

Instâncias c6i

c6i.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1.408	1.408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6i.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c6i.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6i.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c6i.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c6i.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c6i.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c6i.32xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1510m	-Xmx1510m
mapreduce.java.opts	-Xmx3020m	-Xmx3020m
mapreduce.map.memory.mb	1888	1888
mapreduce.reduce.memory.mb	3776	3776
yarn.app.mapreduce.am.resource.mb	3776	3776
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instâncias c6id

c6id.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1.408	1.408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6id.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c6id.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6id.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c6id.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c6id.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c6id.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c6id.32xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1510m	-Xmx1510m
mapreduce.java.opts	-Xmx3020m	-Xmx3020m
mapreduce.map.memory.mb	1888	1888
mapreduce.reduce.memory.mb	3776	3776
yarn.app.mapreduce.am.resource.mb	3776	3776

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instâncias c6in

c6in.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1.408	1.408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6in.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c6in.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6in.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c6in.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c6in.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c6in.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c6in.32xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1510m	-Xmx1510m
mapreduce.java.opts	-Xmx3020m	-Xmx3020m
mapreduce.map.memory.mb	1888	1888
mapreduce.reduce.memory.mb	3776	3776
yarn.app.mapreduce.am.resource.mb	3776	3776
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instâncias c7a

c7a.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1.408	1.408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c7a.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c7a.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c7a.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c7a.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c7a.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c7a.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c7a.32xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1510m	-Xmx1510m
mapreduce.java.opts	-Xmx3020m	-Xmx3020m
mapreduce.map.memory.mb	1888	1888
mapreduce.reduce.memory.mb	3776	3776
yarn.app.mapreduce.am.resource.mb	3776	3776
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

c7a.48xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1527m	-Xmx1527m
mapreduce.java.opts	-Xmx3054m	-Xmx3054m
mapreduce.map.memory.mb	1909	1909
mapreduce.reduce.memory.mb	3818	3818
yarn.app.mapreduce.am.resource.mb	3818	3818
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30608
yarn.nodemanager.resource.memory-mb	366592	30608

Instâncias c7g

c7g.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1.408	1.408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c7g.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c7g.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c7g.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c7g.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c7g.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

Instâncias c7gd

c7gd.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1.408	1.408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c7gd.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c7gd.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c7gd.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c7gd.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c7gd.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

Instâncias c7gn

c7gn.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1.408	1.408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c7gn.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c7gn.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c7gn.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c7gn.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c7gn.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

Instâncias C7i

c7i.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1.408	1.408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c7i.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c7i.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c7i.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c7i.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c7i.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c7i.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c7i.48xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1527m	-Xmx1527m
mapreduce.java.opts	-Xmx3054m	-Xmx3054m
mapreduce.map.memory.mb	1909	1909
mapreduce.reduce.memory.mb	3818	3818
yarn.app.mapreduce.am.resource.mb	3818	3818
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30608
yarn.nodemanager.resource.memory-mb	366592	30608

instâncias c7i-flex

c7i-flex.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1.408	1.408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c7i-flex.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c7i-flex.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c7i-flex.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c7i-flex.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c7i-flex.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

instâncias c8g

c8g.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1.408	1.408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c8g.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c8g.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c8g.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c8g.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c8g.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c8g.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c8g.48xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1527m	-Xmx1527m
mapreduce.java.opts	-Xmx3054m	-Xmx3054m
mapreduce.map.memory.mb	1909	1909
mapreduce.reduce.memory.mb	3818	3818
yarn.app.mapreduce.am.resource.mb	3818	3818

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30608
yarn.nodemanager.resource.memory-mb	366592	30608

instâncias c8gd

c8gd.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1.408	1.408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c8gd.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c8gd.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c8gd.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c8gd.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c8gd.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c8gd.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c8gd.48xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1527m	-Xmx1527m
mapreduce.java.opts	-Xmx3054m	-Xmx3054m
mapreduce.map.memory.mb	1909	1909
mapreduce.reduce.memory.mb	3818	3818
yarn.app.mapreduce.am.resource.mb	3818	3818
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30608
yarn.nodemanager.resource.memory-mb	366592	30608

Instâncias d2

d2.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

d2.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

d2.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

d2.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2417m	-Xmx2417m
mapreduce.java.opts	-Xmx4834m	-Xmx4834m
mapreduce.map.memory.mb	3021	3021
mapreduce.reduce.memory.mb	6042	6042
yarn.app.mapreduce.am.resource.mb	6042	6042

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30194
yarn.nodemanager.resource.memory-mb	241664	30194

Instâncias d3

d3.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

d3.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

d3.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

d3.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instâncias d3en

d3en.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

d3en.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

d3en.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

d3en.6xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.java.opts	-Xmx5700m	-Xmx5700m
mapreduce.map.memory.mb	3563	3563
mapreduce.reduce.memory.mb	7126	7126
yarn.app.mapreduce.am.resource.mb	7126	7126
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	28496
yarn.nodemanager.resource.memory-mb	85504	28496

d3en.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

d3en.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

instâncias f2

f2.6xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-XMX8055m	-XMX8055m
mapreduce.java.opts	-XMX16110m	-XMX16110m
mapreduce.map.memory.mb	10069	10069
mapreduce.reduce.memory.mb	20138	20138
yarn.app.mapreduce.am.resource.mb	20138	20138
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	20146
yarn.nodemanager.resource.memory-mb	241664	20146

f2.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-XMX8192m	-XMX8192m
mapreduce.java.opts	-XMX16384M	-XMX16384M
mapreduce.map.memory.mb	10240	10240
mapreduce.reduce.memory.mb	20480	20480
yarn.app.mapreduce.am.resource.mb	20480	20480

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	20480
yarn.nodemanager.resource.memory-mb	491520	20480

f2.48xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx8294m	-Xmx8294m
mapreduce.java.opts	-Xmx16588m	-Xmx16588m
mapreduce.map.memory.mb	10368	10368
mapreduce.reduce.memory.mb	20736	20736
yarn.app.mapreduce.am.resource.mb	20736	20736
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1990-656	20736
yarn.nodemanager.resource.memory-mb	1990-656	20736

Instâncias g3

g3.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

g3.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

g3.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instâncias g3s

g3s.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

Instâncias do g4dn

g4dn.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144

Opção de configuração	Valor padrão	Com HBase instalado
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

g4dn.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

g4dn.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2867m	-Xmx2867m
mapreduce.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.map.memory.mb	3584	3584
mapreduce.reduce.memory.mb	7168	7168
yarn.app.mapreduce.am.resource.mb	7168	7168
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

g4dn.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

g4dn.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.java.opts	-Xmx6280m	-Xmx6280m
mapreduce.map.memory.mb	3925	3925
mapreduce.reduce.memory.mb	7850	7850
yarn.app.mapreduce.am.resource.mb	7850	7850
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31416
yarn.nodemanager.resource.memory-mb	188416	31416

g4dn.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3174m	-Xmx3174m
mapreduce.java.opts	-Xmx6348m	-Xmx6348m
mapreduce.map.memory.mb	3968	3968
mapreduce.reduce.memory.mb	7936	7936
yarn.app.mapreduce.am.resource.mb	7936	7936
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

Instâncias g5

g5.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

g5.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

g5.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

g5.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

g5.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

g5.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

g5.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

g5.48xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3089m	-Xmx3089m
mapreduce.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.map.memory.mb	3861	3861
mapreduce.reduce.memory.mb	7722	7722
yarn.app.mapreduce.am.resource.mb	7722	7722
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30952
yarn.nodemanager.resource.memory-mb	741376	30952

instâncias g6

g6.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.java.opts	-Xmx4608m	-Xmx4608m
mapreduce.map.memory.mb	2.880	2.880
mapreduce.reduce.memory.mb	5760	5760
yarn.app.mapreduce.am.resource.mb	5760	5760
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11520	5760
yarn.nodemanager.resource.memory-mb	11520	5760

g6.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.java.opts	-Xmx4608m	-Xmx4608m
mapreduce.map.memory.mb	2.880	2.880
mapreduce.reduce.memory.mb	5760	5760
yarn.app.mapreduce.am.resource.mb	5760	5760

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23040	11520
yarn.nodemanager.resource.memory-mb	23040	11520

g6.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2662m	-Xmx2662m
mapreduce.java.opts	-XMX5324m	-XMX5324m
mapreduce.map.memory.mb	3328	3328
mapreduce.reduce.memory.mb	6656	6656
yarn.app.mapreduce.am.resource.mb	6656	6656
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	53248	26624
yarn.nodemanager.resource.memory-mb	53248	26624

g6.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-XMX2893m	-XMX2893m
mapreduce.java.opts	-XMX5786m	-XMX5786m
mapreduce.map.memory.mb	3616	3616
mapreduce.reduce.memory.mb	7232	7232
yarn.app.mapreduce.am.resource.mb	7232	7232
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	115712	28928
yarn.nodemanager.resource.memory-mb	115712	28928

g6.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-XMX2953m	-XMX2953m
mapreduce.java.opts	-XMX5906m	-XMX5906m
mapreduce.map.memory.mb	3691	3691
mapreduce.reduce.memory.mb	7382	7382
yarn.app.mapreduce.am.resource.mb	7382	7382

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	177152	29512
yarn.nodemanager.resource.memory-mb	177152	29512

g6.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-XMX2995m	-XMX2995m
mapreduce.java.opts	-XMX5990m	-XMX5990m
mapreduce.map.memory.mb	3744	3744
mapreduce.reduce.memory.mb	7488	7488
yarn.app.mapreduce.am.resource.mb	7488	7488
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	239616	29952
yarn.nodemanager.resource.memory-mb	239616	29952

g6.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-XMX3038m	-XMX3038m
mapreduce.java.opts	-XMX6076m	-XMX6076m
mapreduce.map.memory.mb	3797	3797
mapreduce.reduce.memory.mb	7594	7594
yarn.app.mapreduce.am.resource.mb	7594	7594
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	364544	30408
yarn.nodemanager.resource.memory-mb	364544	30408

g6.48xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	737280	30720
yarn.nodemanager.resource.memory-mb	737280	30720

instâncias g6e

g6e.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

g6e.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

g6e.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

g6e.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

g6e.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

g6e.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

g6e.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

g6e.48xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	1530	1530
yarn.app.mapreduce.am.resource.mb	1530	1530
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

instâncias gr6

gr6.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-XMX5786m	-XMX5786m
mapreduce.java.opts	-XMX11572m	-XMX11572m
mapreduce.map.memory.mb	7232	7232
mapreduce.reduce.memory.mb	1464	1464
yarn.app.mapreduce.am.resource.mb	1464	1464

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	115712	28928
yarn.nodemanager.resource.memory-mb	115712	28928

gr6.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-XMX5990m	-XMX5990m
mapreduce.java.opts	- XMX11980m	- XMX11980m
mapreduce.map.memory.mb	7488	7488
mapreduce.reduce.memory.mb	14976	14976
yarn.app.mapreduce.am.resource.mb	14976	14976
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	239616	29952
yarn.nodemanager.resource.memory-mb	239616	29952

Instâncias h1

h1.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

h1.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2867m	-Xmx2867m
mapreduce.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.map.memory.mb	3584	3584
mapreduce.reduce.memory.mb	7168	7168
yarn.app.mapreduce.am.resource.mb	7168	7168

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

h1.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

h1.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3174m	-Xmx3174m
mapreduce.java.opts	-Xmx6348m	-Xmx6348m
mapreduce.map.memory.mb	3968	3968
mapreduce.reduce.memory.mb	7936	7936
yarn.app.mapreduce.am.resource.mb	7936	7936
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

Instâncias i2

i2.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

i2.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

i2.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

i2.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instâncias i3

i3.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

i3.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

i3.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

i3.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

i3.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instâncias i3en

i3en.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4915m	-Xmx4915m
mapreduce.java.opts	-Xmx9830m	-Xmx9830m
mapreduce.map.memory.mb	6144	6144
mapreduce.reduce.memory.mb	12288	12288
yarn.app.mapreduce.am.resource.mb	12288	12288

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

i3en.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.java.opts	-Xmx11468m	-Xmx11468m
mapreduce.map.memory.mb	7168	7168
mapreduce.reduce.memory.mb	14336	14336
yarn.app.mapreduce.am.resource.mb	14336	14336
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

i3en.3xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6007m	-Xmx6007m
mapreduce.java.opts	-Xmx12014m	-Xmx12014m
mapreduce.map.memory.mb	7509	7509
mapreduce.reduce.memory.mb	15018	15018
yarn.app.mapreduce.am.resource.mb	15018	15018
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	90112	30040
yarn.nodemanager.resource.memory-mb	90112	30040

i3en.6xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6281m	-Xmx6281m
mapreduce.java.opts	-Xmx12562m	-Xmx12562m
mapreduce.map.memory.mb	7851	7851
mapreduce.reduce.memory.mb	15702	15702
yarn.app.mapreduce.am.resource.mb	15702	15702

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31396
yarn.nodemanager.resource.memory-mb	188416	31396

i3en.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6417m	-Xmx6417m
mapreduce.java.opts	-Xmx12834m	-Xmx12834m
mapreduce.map.memory.mb	8021	8021
mapreduce.reduce.memory.mb	16042	16042
yarn.app.mapreduce.am.resource.mb	16042	16042
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32100
yarn.nodemanager.resource.memory-mb	385024	32100

i3en.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6486m	-Xmx6486m
mapreduce.java.opts	-Xmx12972m	-Xmx12972m
mapreduce.map.memory.mb	8107	8107
mapreduce.reduce.memory.mb	16214	16214
yarn.app.mapreduce.am.resource.mb	16214	16214
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	778240	32396
yarn.nodemanager.resource.memory-mb	778240	32396

Instâncias i4g

i4g.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

i4g.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

i4g.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

i4g.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

i4g.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instâncias i4i

i4i.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

i4i.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

i4i.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

i4i.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

i4i.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

i4i.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

i4i.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

i4i.32xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	774	774
mapreduce.reduce.memory.mb	15488	15488
yarn.app.mapreduce.am.resource.mb	15488	15488

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

instâncias i7i

i7i.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

i7i.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

i7i.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

i7i.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

i7i.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

i7i.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

i7i.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

i7i.48xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	1530	1530
yarn.app.mapreduce.am.resource.mb	1530	1530
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

instâncias i7ie

i7ie.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

i7ie.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

i7ie.3xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5700m	-Xmx5700m
mapreduce.java.opts	-XMX 11400m	-XMX 11400m
mapreduce.map.memory.mb	7125	7125
mapreduce.reduce.memory.mb	14250	14250
yarn.app.mapreduce.am.resource.mb	14250	14250
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	28504
yarn.nodemanager.resource.memory-mb	85504	28504

i7ie.6xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-XMX5974m	-XMX5974m
mapreduce.java.opts	-XMX11948m	-XMX11948m
mapreduce.map.memory.mb	7467	7467
mapreduce.reduce.memory.mb	14934	14934
yarn.app.mapreduce.am.resource.mb	14934	14934

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29860
yarn.nodemanager.resource.memory-mb	179200	29860

i7ie.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

i7ie.18xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-XMX6155m	-XMX6155m
mapreduce.java.opts	-XMX12310m	-XMX12310m
mapreduce.map.memory.mb	7694	7694
mapreduce.reduce.memory.mb	1538	1538
yarn.app.mapreduce.am.resource.mb	1538	1538
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	553984	30792
yarn.nodemanager.resource.memory-mb	553984	30792

i7ie.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

i7ie.48xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	1530	1530
yarn.app.mapreduce.am.resource.mb	1530	1530
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

instâncias i8g

i8g.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

i8g.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

i8g.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

i8g.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

i8g.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

i8g.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

i8g.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

i8g.48xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	1530	1530
yarn.app.mapreduce.am.resource.mb	1530	1530

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

Instâncias im4gn

im4gn.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

im4gn.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

im4gn.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

im4gn.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

im4gn.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instâncias is4gen

is4gen.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3514m	-Xmx3514m
mapreduce.java.opts	-Xmx7028m	-Xmx7028m
mapreduce.map.memory.mb	4393	4393
mapreduce.reduce.memory.mb	8786	8786
yarn.app.mapreduce.am.resource.mb	8786	8786

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	17572,12	8.786,06
yarn.nodemanager.resource.memory-mb	17572,12	8.786,06

is4gen.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3866m	-Xmx3866m
mapreduce.java.opts	-Xmx7732m	-Xmx7732m
mapreduce.map.memory.mb	4832	4832
mapreduce.reduce.memory.mb	9664	9664
yarn.app.mapreduce.am.resource.mb	9664	9664
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	38656	19328
yarn.nodemanager.resource.memory-mb	38656	19328

is4gen.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4275m	-Xmx4275m
mapreduce.java.opts	-Xmx8550m	-Xmx8550m
mapreduce.map.memory.mb	5344	5344
mapreduce.reduce.memory.mb	1068	1068
yarn.app.mapreduce.am.resource.mb	1068	1068
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32064
yarn.nodemanager.resource.memory-mb	85504	32064

is4gen.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4480m	-Xmx4480m
mapreduce.java.opts	-Xmx8960m	-Xmx8960m
mapreduce.map.memory.mb	5600	5600
mapreduce.reduce.memory.mb	11200	11200
yarn.app.mapreduce.am.resource.mb	11200	11200

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	2400
yarn.nodemanager.resource.memory-mb	179200	2400

Instâncias m1

m1.small

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx288m	-Xmx288m
mapreduce.java.opts	-Xmx288m	-Xmx288m
mapreduce.map.memory.mb	512	512
mapreduce.reduce.memory.mb	512	512
yarn.app.mapreduce.am.resource.mb	512	512
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	512	512
yarn.nodemanager.resource.memory-mb	1024	512

m1.medium

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx512m	-Xmx512m
mapreduce.java.opts	-Xmx768m	-Xmx768m
mapreduce.map.memory.mb	768	768
mapreduce.reduce.memory.mb	1024	1024
yarn.app.mapreduce.am.resource.mb	1024	1024
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	2048	1024
yarn.nodemanager.resource.memory-mb	2048	1024

m1.large

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx512m	-Xmx512m
mapreduce.java.opts	-Xmx1024m	-Xmx1024m
mapreduce.map.memory.mb	768	768
mapreduce.reduce.memory.mb	1536	1536
yarn.app.mapreduce.am.resource.mb	1536	1536

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	3072	2560
yarn.nodemanager.resource.memory-mb	5120	2560

m1.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx512m	-Xmx512m
mapreduce.java.opts	-Xmx1536m	-Xmx1536m
mapreduce.map.memory.mb	768	768
mapreduce.reduce.memory.mb	2048	2048
yarn.app.mapreduce.am.resource.mb	2048	2048
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	8192	6144
yarn.nodemanager.resource.memory-mb	12288	6144

Instâncias m2

m2.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx864m	-Xmx864m
mapreduce.java.opts	-Xmx1536m	-Xmx1536m
mapreduce.map.memory.mb	1024	1024
mapreduce.reduce.memory.mb	2048	2048
yarn.app.mapreduce.am.resource.mb	2048	2048
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	7168	7168
yarn.nodemanager.resource.memory-mb	14336	7168

m2.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1280m	-Xmx1280m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	2560	2560
yarn.app.mapreduce.am.resource.mb	2560	2560

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	8192	15360
yarn.nodemanager.resource.memory-mb	30720	15360

m2.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1280m	-Xmx1280m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	2560	2560
yarn.app.mapreduce.am.resource.mb	2560	2560
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	8192	30720
yarn.nodemanager.resource.memory-mb	61440	30720

Instâncias m3

m3.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1152m	-Xmx1152m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1440	1440
mapreduce.reduce.memory.mb	2.880	2.880
yarn.app.mapreduce.am.resource.mb	2.880	2.880
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11520	5760
yarn.nodemanager.resource.memory-mb	11520	5760

m3.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1152m	-Xmx1152m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1440	1440
mapreduce.reduce.memory.mb	2.880	2.880
yarn.app.mapreduce.am.resource.mb	2.880	2.880

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23040	11520
yarn.nodemanager.resource.memory-mb	23040	11520

Instâncias m4

m4.large

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	6144	3072
yarn.nodemanager.resource.memory-mb	6144	3072

m4.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

m4.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

m4.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1434m	-Xmx1434m
mapreduce.java.opts	-Xmx2868m	-Xmx2868m
mapreduce.map.memory.mb	1792	1792
mapreduce.reduce.memory.mb	3584	3584
yarn.app.mapreduce.am.resource.mb	3584	3584
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

m4.10xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1557m	-Xmx1557m
mapreduce.java.opts	-Xmx3114m	-Xmx3114m
mapreduce.map.memory.mb	1946	1946
mapreduce.reduce.memory.mb	3892	3892
yarn.app.mapreduce.am.resource.mb	3892	3892
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	155648	31104
yarn.nodemanager.resource.memory-mb	155648	31104

m4.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx1587m	-Xmx1587m
mapreduce.java.opts	-Xmx3174m	-Xmx3174m
mapreduce.map.memory.mb	1984	1984
mapreduce.reduce.memory.mb	3968	3968
yarn.app.mapreduce.am.resource.mb	3968	3968

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

Instâncias m5

m5.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

m5.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

m5.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2867m	-Xmx2867m
mapreduce.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.map.memory.mb	3584	3584
mapreduce.reduce.memory.mb	7168	7168
yarn.app.mapreduce.am.resource.mb	7168	7168

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

m5.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

m5.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.java.opts	-Xmx6280m	-Xmx6280m
mapreduce.map.memory.mb	3925	3925
mapreduce.reduce.memory.mb	7850	7850
yarn.app.mapreduce.am.resource.mb	7850	7850
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31416
yarn.nodemanager.resource.memory-mb	188416	31416

m5.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3174m	-Xmx3174m
mapreduce.java.opts	-Xmx6348m	-Xmx6348m
mapreduce.map.memory.mb	3968	3968
mapreduce.reduce.memory.mb	7936	7936
yarn.app.mapreduce.am.resource.mb	7936	7936

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

m5.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3209m	-Xmx3209m
mapreduce.java.opts	-Xmx6418m	-Xmx6418m
mapreduce.map.memory.mb	4011	4011
mapreduce.reduce.memory.mb	8022	8022
yarn.app.mapreduce.am.resource.mb	8022	8022
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32056
yarn.nodemanager.resource.memory-mb	385024	32056

Instâncias m5a

m5a.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

m5a.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

m5a.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2867m	-Xmx2867m
mapreduce.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.map.memory.mb	3584	3584
mapreduce.reduce.memory.mb	7168	7168
yarn.app.mapreduce.am.resource.mb	7168	7168
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

m5a.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

m5a.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.java.opts	-Xmx6280m	-Xmx6280m
mapreduce.map.memory.mb	3925	3925
mapreduce.reduce.memory.mb	7850	7850
yarn.app.mapreduce.am.resource.mb	7850	7850

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31416
yarn.nodemanager.resource.memory-mb	188416	31416

m5a.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3174m	-Xmx3174m
mapreduce.java.opts	-Xmx6348m	-Xmx6348m
mapreduce.map.memory.mb	3968	3968
mapreduce.reduce.memory.mb	7936	7936
yarn.app.mapreduce.am.resource.mb	7936	7936
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

m5a.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3209m	-Xmx3209m
mapreduce.java.opts	-Xmx6418m	-Xmx6418m
mapreduce.map.memory.mb	4011	4011
mapreduce.reduce.memory.mb	8022	8022
yarn.app.mapreduce.am.resource.mb	8022	8022
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32056
yarn.nodemanager.resource.memory-mb	385024	32056

Instâncias m5ad

m5ad.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m5ad.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m5ad.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m5ad.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m5ad.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	2980
yarn.nodemanager.resource.memory-mb	179200	2980

m5ad.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m5ad.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

Instâncias m5d

m5d.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

m5d.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

m5d.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2867m	-Xmx2867m
mapreduce.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.map.memory.mb	3584	3584
mapreduce.reduce.memory.mb	7168	7168
yarn.app.mapreduce.am.resource.mb	7168	7168

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

m5d.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

m5d.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.java.opts	-Xmx6280m	-Xmx6280m
mapreduce.map.memory.mb	3925	3925
mapreduce.reduce.memory.mb	7850	7850
yarn.app.mapreduce.am.resource.mb	7850	7850
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31416
yarn.nodemanager.resource.memory-mb	188416	31416

m5d.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3174m	-Xmx3174m
mapreduce.java.opts	-Xmx6348m	-Xmx6348m
mapreduce.map.memory.mb	3968	3968
mapreduce.reduce.memory.mb	7936	7936
yarn.app.mapreduce.am.resource.mb	7936	7936

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

m5d.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3209m	-Xmx3209m
mapreduce.java.opts	-Xmx6418m	-Xmx6418m
mapreduce.map.memory.mb	4011	4011
mapreduce.reduce.memory.mb	8022	8022
yarn.app.mapreduce.am.resource.mb	8022	8022
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32056
yarn.nodemanager.resource.memory-mb	385024	32056

Instâncias m5dn

m5dn.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m5dn.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m5dn.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m5dn.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m5dn.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	2980
yarn.nodemanager.resource.memory-mb	179200	2980

m5dn.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m5dn.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

Instâncias m5n

m5n.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m5n.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m5n.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m5n.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m5n.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	2980
yarn.nodemanager.resource.memory-mb	179200	2980

m5n.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m5n.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

instâncias m5zn

m5zn.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.java.opts	-Xmx4608m	-Xmx4608m
mapreduce.map.memory.mb	2.880	2.880
mapreduce.reduce.memory.mb	5760	5760
yarn.app.mapreduce.am.resource.mb	5760	5760
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11520	5760
yarn.nodemanager.resource.memory-mb	11520	5760

m5zn.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m5zn.3xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2577m	-Xmx2577m
mapreduce.java.opts	-Xmx5154m	-Xmx5154m
mapreduce.map.memory.mb	3221	3221
mapreduce.reduce.memory.mb	6442	6442
yarn.app.mapreduce.am.resource.mb	6442	6442

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	38656	19328
yarn.nodemanager.resource.memory-mb	38656	19328

m5zn.6xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.java.opts	-Xmx5700m	-Xmx5700m
mapreduce.map.memory.mb	3563	3563
mapreduce.reduce.memory.mb	7126	7126
yarn.app.mapreduce.am.resource.mb	7126	7126
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	28496
yarn.nodemanager.resource.memory-mb	85504	28496

m5zn.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	2980
yarn.nodemanager.resource.memory-mb	179200	2980

Instâncias m6a

m6a.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m6a.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m6a.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6a.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6a.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	2980
yarn.nodemanager.resource.memory-mb	179200	2980

m6a.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m6a.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m6a.32xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

m6a.48xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3089m	-Xmx3089m
mapreduce.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.map.memory.mb	3861	3861
mapreduce.reduce.memory.mb	7722	7722
yarn.app.mapreduce.am.resource.mb	7722	7722
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30952
yarn.nodemanager.resource.memory-mb	741376	30952

Instâncias m6g

m6g.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m6g.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m6g.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6g.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6g.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	181248	30208
yarn.nodemanager.resource.memory-mb	181248	30208

m6g.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instâncias m6gd

m6gd.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m6gd.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m6gd.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6gd.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6gd.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	181248	30208
yarn.nodemanager.resource.memory-mb	181248	30208

m6gd.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instâncias m6i

m6i.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m6i.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m6i.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6i.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6i.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	181248	30208
yarn.nodemanager.resource.memory-mb	181248	30208

m6i.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m6i.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m6i.32xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instâncias m6id

m6id.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m6id.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m6id.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6id.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6id.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	2980
yarn.nodemanager.resource.memory-mb	179200	2980

m6id.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m6id.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m6id.32xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instâncias m6idn

m6idn.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m6idn.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m6idn.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6idn.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6idn.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	2980
yarn.nodemanager.resource.memory-mb	179200	2980

m6idn.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m6idn.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m6idn.32xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instâncias m6in

m6in.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m6in.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m6in.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6in.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6in.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	181248	30208
yarn.nodemanager.resource.memory-mb	181248	30208

m6in.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m6in.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m6in.32xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instâncias M7a

m7a.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m7a.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m7a.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m7a.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m7a.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	2980
yarn.nodemanager.resource.memory-mb	179200	2980

m7a.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m7a.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m7a.32xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

m7a.48xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3089m	-Xmx3089m
mapreduce.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.map.memory.mb	3861	3861
mapreduce.reduce.memory.mb	7722	7722
yarn.app.mapreduce.am.resource.mb	7722	7722
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30952
yarn.nodemanager.resource.memory-mb	741376	30952

Instâncias m7g

m7g.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m7g.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m7g.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m7g.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m7g.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	2980
yarn.nodemanager.resource.memory-mb	179200	2980

m7g.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instâncias m7gd

m7gd.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m7gd.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m7gd.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m7gd.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m7gd.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	2980
yarn.nodemanager.resource.memory-mb	179200	2980

m7gd.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instâncias M7i

m7i.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m7i.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m7i.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m7i.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m7i.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	181248	30208
yarn.nodemanager.resource.memory-mb	181248	30208

m7i.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m7i.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m7i.48xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3089m	-Xmx3089m
mapreduce.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.map.memory.mb	3861	3861
mapreduce.reduce.memory.mb	7722	7722
yarn.app.mapreduce.am.resource.mb	7722	7722

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30952
yarn.nodemanager.resource.memory-mb	741376	30952

Instâncias M7i-flex

m7i-flex.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m7i-flex.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m7i-flex.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m7i-flex.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m7i-flex.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	2980
yarn.nodemanager.resource.memory-mb	179200	2980

m7i-flex.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

instâncias m8g

m8g.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m8g.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m8g.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m8g.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m8g.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	181248	30208
yarn.nodemanager.resource.memory-mb	181248	30208

m8g.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m8g.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m8g.48xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3089m	-Xmx3089m
mapreduce.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.map.memory.mb	3861	3861
mapreduce.reduce.memory.mb	7722	7722
yarn.app.mapreduce.am.resource.mb	7722	7722
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30952
yarn.nodemanager.resource.memory-mb	741376	30952

instâncias m8gd

m8gd.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m8gd.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m8gd.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m8gd.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m8gd.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	2980
yarn.nodemanager.resource.memory-mb	179200	2980

m8gd.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m8gd.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m8gd.48xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3089m	-Xmx3089m
mapreduce.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.map.memory.mb	3861	3861
mapreduce.reduce.memory.mb	7722	7722
yarn.app.mapreduce.am.resource.mb	7722	7722
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30952
yarn.nodemanager.resource.memory-mb	741376	30952

Instâncias p2

p2.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.java.opts	-Xmx21708m	-Xmx21708m
mapreduce.map.memory.mb	13568	13568
mapreduce.reduce.memory.mb	27136	27136
yarn.app.mapreduce.am.resource.mb	27136	27136
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

p2.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.java.opts	-Xmx24576m	-Xmx24576m
mapreduce.map.memory.mb	15360	15360
mapreduce.reduce.memory.mb	30720	30720
yarn.app.mapreduce.am.resource.mb	30720	30720

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

p2.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx9267m	-Xmx9267m
mapreduce.java.opts	-Xmx18534m	-Xmx18534m
mapreduce.map.memory.mb	11584	11584
mapreduce.reduce.memory.mb	23168	23168
yarn.app.mapreduce.am.resource.mb	23168	23168
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	23168
yarn.nodemanager.resource.memory-mb	741376	23168

Instâncias p3

p3.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

p3.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

p3.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

instâncias p4d

p4d.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-XMX9302m	-XMX9302m
mapreduce.java.opts	-XMX18604M	-XMX18604M
mapreduce.map.memory.mb	11627	11627
mapreduce.reduce.memory.mb	23254	23254
yarn.app.mapreduce.am.resource.mb	23254	23254
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1116160	23222
yarn.nodemanager.resource.memory-mb	1116160	23222

Instâncias p5

p5.48xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx8294m	-Xmx8294m
mapreduce.java.opts	-Xmx16588m	-Xmx16588m
mapreduce.map.memory.mb	10368	10368
mapreduce.reduce.memory.mb	20736	20736

Opção de configuração	Valor padrão	Com HBase instalado
yarn.app.mapreduce.am.resource.mb	20736	20736
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1990-656	20736
yarn.nodemanager.resource.memory-mb	1990-656	20736

Instâncias r3

r3.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r3.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r3.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r3.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instâncias r4

r4.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r4.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r4.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r4.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r4.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instâncias r5

r5.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4915m	-Xmx4915m
mapreduce.java.opts	-Xmx9830m	-Xmx9830m
mapreduce.map.memory.mb	6144	6144
mapreduce.reduce.memory.mb	12288	12288
yarn.app.mapreduce.am.resource.mb	12288	12288
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

r5.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.java.opts	-Xmx11468m	-Xmx11468m
mapreduce.map.memory.mb	7168	7168
mapreduce.reduce.memory.mb	14336	14336
yarn.app.mapreduce.am.resource.mb	14336	14336
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

r5.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

r5.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6349m	-Xmx6349m
mapreduce.java.opts	-Xmx12698m	-Xmx12698m
mapreduce.map.memory.mb	7936	7936
mapreduce.reduce.memory.mb	15872	15872
yarn.app.mapreduce.am.resource.mb	15872	15872
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

r5.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6417m	-Xmx6417m
mapreduce.java.opts	-Xmx12834m	-Xmx12834m
mapreduce.map.memory.mb	8021	8021
mapreduce.reduce.memory.mb	16042	16042
yarn.app.mapreduce.am.resource.mb	16042	16042
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32100
yarn.nodemanager.resource.memory-mb	385024	32100

r5.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6451m	-Xmx6451m
mapreduce.java.opts	-Xmx12902m	-Xmx12902m
mapreduce.map.memory.mb	8064	8064
mapreduce.reduce.memory.mb	16128	16128
yarn.app.mapreduce.am.resource.mb	16128	16128

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	516096	32256
yarn.nodemanager.resource.memory-mb	516096	32256

r5.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6486m	-Xmx6486m
mapreduce.java.opts	-Xmx12972m	-Xmx12972m
mapreduce.map.memory.mb	8107	8107
mapreduce.reduce.memory.mb	16214	16214
yarn.app.mapreduce.am.resource.mb	16214	16214
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	778240	32396
yarn.nodemanager.resource.memory-mb	778240	32396

Instâncias r5a

r5a.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4915m	-Xmx4915m
mapreduce.java.opts	-Xmx9830m	-Xmx9830m
mapreduce.map.memory.mb	6144	6144
mapreduce.reduce.memory.mb	12288	12288
yarn.app.mapreduce.am.resource.mb	12288	12288
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

r5a.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.java.opts	-Xmx11468m	-Xmx11468m
mapreduce.map.memory.mb	7168	7168
mapreduce.reduce.memory.mb	14336	14336
yarn.app.mapreduce.am.resource.mb	14336	14336

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

r5a.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

r5a.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6349m	-Xmx6349m
mapreduce.java.opts	-Xmx12698m	-Xmx12698m
mapreduce.map.memory.mb	7936	7936
mapreduce.reduce.memory.mb	15872	15872
yarn.app.mapreduce.am.resource.mb	15872	15872
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

r5a.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6417m	-Xmx6417m
mapreduce.java.opts	-Xmx12834m	-Xmx12834m
mapreduce.map.memory.mb	8021	8021
mapreduce.reduce.memory.mb	16042	16042
yarn.app.mapreduce.am.resource.mb	16042	16042

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32100
yarn.nodemanager.resource.memory-mb	385024	32100

r5a.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6451m	-Xmx6451m
mapreduce.java.opts	-Xmx12902m	-Xmx12902m
mapreduce.map.memory.mb	8064	8064
mapreduce.reduce.memory.mb	16128	16128
yarn.app.mapreduce.am.resource.mb	16128	16128
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	516096	32256
yarn.nodemanager.resource.memory-mb	516096	32256

r5a.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6486m	-Xmx6486m
mapreduce.java.opts	-Xmx12972m	-Xmx12972m
mapreduce.map.memory.mb	8107	8107
mapreduce.reduce.memory.mb	16214	16214
yarn.app.mapreduce.am.resource.mb	16214	16214
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	778240	32396
yarn.nodemanager.resource.memory-mb	778240	32396

Instâncias r5ad

r5ad.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r5ad.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r5ad.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r5ad.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r5ad.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r5ad.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6246m	-Xmx6246m
mapreduce.java.opts	-Xmx12492m	-Xmx12492m
mapreduce.map.memory.mb	7808	7808
mapreduce.reduce.memory.mb	15616	15616
yarn.app.mapreduce.am.resource.mb	15616	15616
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	499712	31232
yarn.nodemanager.resource.memory-mb	499712	31232

r5ad.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

Instâncias r5b

r5b.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r5n.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r5b.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r5b.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r5b.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r5b.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r5b.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

Instâncias r5d

r5d.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4915m	-Xmx4915m
mapreduce.java.opts	-Xmx9830m	-Xmx9830m
mapreduce.map.memory.mb	6144	6144
mapreduce.reduce.memory.mb	12288	12288
yarn.app.mapreduce.am.resource.mb	12288	12288
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

r5d.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.java.opts	-Xmx11468m	-Xmx11468m
mapreduce.map.memory.mb	7168	7168
mapreduce.reduce.memory.mb	14336	14336
yarn.app.mapreduce.am.resource.mb	14336	14336

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

r5d.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

r5d.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6349m	-Xmx6349m
mapreduce.java.opts	-Xmx12698m	-Xmx12698m
mapreduce.map.memory.mb	7936	7936
mapreduce.reduce.memory.mb	15872	15872
yarn.app.mapreduce.am.resource.mb	15872	15872
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

r5d.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6417m	-Xmx6417m
mapreduce.java.opts	-Xmx12834m	-Xmx12834m
mapreduce.map.memory.mb	8021	8021
mapreduce.reduce.memory.mb	16042	16042
yarn.app.mapreduce.am.resource.mb	16042	16042

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32100
yarn.nodemanager.resource.memory-mb	385024	32100

r5d.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6451m	-Xmx6451m
mapreduce.java.opts	-Xmx12902m	-Xmx12902m
mapreduce.map.memory.mb	8064	8064
mapreduce.reduce.memory.mb	16128	16128
yarn.app.mapreduce.am.resource.mb	16128	16128
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	516096	32256
yarn.nodemanager.resource.memory-mb	516096	32256

r5d.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6486m	-Xmx6486m
mapreduce.java.opts	-Xmx12972m	-Xmx12972m
mapreduce.map.memory.mb	8107	8107
mapreduce.reduce.memory.mb	16214	16214
yarn.app.mapreduce.am.resource.mb	16214	16214
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	778240	32396
yarn.nodemanager.resource.memory-mb	778240	32396

Instâncias r5dn

r5dn.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r5dn.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r5dn.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r5dn.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r5dn.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r5dn.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r5dn.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

Instâncias r5n

r5n.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r5n.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r5n.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r5n.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r5n.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r5n.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r5n.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

Instâncias r6a

r6a.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6a.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r6a.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r6a.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r6a.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r6a.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r6a.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r6a.32xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	774	774
mapreduce.reduce.memory.mb	15488	15488
yarn.app.mapreduce.am.resource.mb	15488	15488

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

r6a.48xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	1530	1530
yarn.app.mapreduce.am.resource.mb	1530	1530
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

Instâncias r6g

r6g.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6g.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r6g.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r6g.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r6g.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r6g.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instâncias r6gd

r6gd.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6gd.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r6gd.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r6gd.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r6gd.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r6gd.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instâncias r6i

r6i.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6i.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r6i.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r6i.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r6i.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r6i.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r6i.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r6i.32xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6029m	-Xmx6029m
mapreduce.java.opts	-Xmx12058m	-Xmx12058m
mapreduce.map.memory.mb	7536	7536
mapreduce.reduce.memory.mb	15072	15072
yarn.app.mapreduce.am.resource.mb	15072	15072

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	964608	3014
yarn.nodemanager.resource.memory-mb	964608	3014

Instâncias r6iD

r6id.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6id.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r6id.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r6id.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r6id.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r6id.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r6id.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r6id.32xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	774	774
mapreduce.reduce.memory.mb	1548	1548
yarn.app.mapreduce.am.resource.mb	1548	1548
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

instâncias r6idn

r6idn.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6idn.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r6idn.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r6idn.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r6idn.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r6idn.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r6idn.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r6idn.32xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	774	774
mapreduce.reduce.memory.mb	15488	15488
yarn.app.mapreduce.am.resource.mb	15488	15488
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

Instâncias r6in

r6in.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6in.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r6in.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r6in.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r6in.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r6in.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r6in.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r6in.32xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	774	774
mapreduce.reduce.memory.mb	15488	15488
yarn.app.mapreduce.am.resource.mb	15488	15488

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

Instâncias R7a

r7a.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r7a.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r7a.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r7a.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r7a.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r7a.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r7a.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r7a.32xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	774	774
mapreduce.reduce.memory.mb	15488	15488
yarn.app.mapreduce.am.resource.mb	15488	15488
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

r7a.48xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	1530	1530
yarn.app.mapreduce.am.resource.mb	1530	1530

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

Instâncias r7g

r7g.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r7g.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r7g.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r7g.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r7g.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r7g.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instâncias r7gd

r7gd.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r7gd.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r7gd.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r7gd.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r7gd.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r7gd.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instâncias r7i

r7i.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r7i.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r7i.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r7i.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r7i.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r7i.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r7i.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r7i.48xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	1530	1530
yarn.app.mapreduce.am.resource.mb	1530	1530
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

Instâncias R7iz

r7iz.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r7iz.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r7iz.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r7iz.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r7iz.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r7iz.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r7iz.32xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6029m	-Xmx6029m
mapreduce.java.opts	-Xmx12058m	-Xmx12058m
mapreduce.map.memory.mb	7536	7536
mapreduce.reduce.memory.mb	15072	15072
yarn.app.mapreduce.am.resource.mb	15072	15072

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	964608	3014
yarn.nodemanager.resource.memory-mb	964608	3014

instâncias r8g

r8g.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r8g.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r8g.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r8g.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r8g.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r8g.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r8g.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r8g.48xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	1530	1530
yarn.app.mapreduce.am.resource.mb	1530	1530
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

instâncias r8gd

r8gd.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r8gd.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r8gd.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r8gd.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r8gd.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r8gd.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r8gd.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r8gd.48xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	1530	1530
yarn.app.mapreduce.am.resource.mb	1530	1530
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

Instâncias x1

x1.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx12058m	-Xmx12058m
mapreduce.java.opts	-Xmx24116m	-Xmx24116m
mapreduce.map.memory.mb	15072	15072
mapreduce.reduce.memory.mb	3014	3014
yarn.app.mapreduce.am.resource.mb	3014	3014
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	964608	3014
yarn.nodemanager.resource.memory-mb	964608	3014

x1.32xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx12109m	-Xmx12109m
mapreduce.java.opts	-Xmx24218m	-Xmx24218m
mapreduce.map.memory.mb	15136	15136
mapreduce.reduce.memory.mb	30272	30272
yarn.app.mapreduce.am.resource.mb	30272	30272

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1937-408	30272
yarn.nodemanager.resource.memory-mb	1937-408	30272

Instâncias x1e

x1e.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx22682m	-Xmx22682m
mapreduce.java.opts	-Xmx45364m	-Xmx45364m
mapreduce.map.memory.mb	28352	28352
mapreduce.reduce.memory.mb	56704	56704
yarn.app.mapreduce.am.resource.mb	56704	56704
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	113408	0
yarn.nodemanager.resource.memory-mb	113408	0

x1e.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx23501m	-Xmx23501m
mapreduce.java.opts	-Xmx47002m	-Xmx47002m
mapreduce.map.memory.mb	29376	29376
mapreduce.reduce.memory.mb	58752	58752
yarn.app.mapreduce.am.resource.mb	58752	58752
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	235008	0
yarn.nodemanager.resource.memory-mb	235008	0

x1e.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx23910m	-Xmx23910m
mapreduce.java.opts	-Xmx47820m	-Xmx47820m
mapreduce.map.memory.mb	2988	2988
mapreduce.reduce.memory.mb	59776	59776
yarn.app.mapreduce.am.resource.mb	59776	59776

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	478208	0
yarn.nodemanager.resource.memory-mb	478208	0

x1e.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx24115m	-Xmx24115m
mapreduce.java.opts	-Xmx48230m	-Xmx48230m
mapreduce.map.memory.mb	3014	3014
mapreduce.reduce.memory.mb	60288	60288
yarn.app.mapreduce.am.resource.mb	60288	60288
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	964608	0
yarn.nodemanager.resource.memory-mb	964608	0

x1e.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx24218m	-Xmx24218m
mapreduce.java.opts	-Xmx48436m	-Xmx48436m
mapreduce.map.memory.mb	30272	30272
mapreduce.reduce.memory.mb	60544	60544
yarn.app.mapreduce.am.resource.mb	60544	60544
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1937-408	0
yarn.nodemanager.resource.memory-mb	1937-408	0

x1e.32xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx24269m	-Xmx24269m
mapreduce.java.opts	-Xmx48538m	-Xmx48538m
mapreduce.map.memory.mb	30336	30336
mapreduce.reduce.memory.mb	60672	60672
yarn.app.mapreduce.am.resource.mb	60672	60672

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	3883008	0
yarn.nodemanager.resource.memory-mb	3883008	0

Instâncias x2gd

x2gd.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.java.opts	-Xmx21708m	-Xmx21708m
mapreduce.map.memory.mb	13568	13568
mapreduce.reduce.memory.mb	27136	27136
yarn.app.mapreduce.am.resource.mb	27136	27136
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

x2gd.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.java.opts	-Xmx23348m	-Xmx23348m
mapreduce.map.memory.mb	14592	14592
mapreduce.reduce.memory.mb	29184	29184
yarn.app.mapreduce.am.resource.mb	29184	29184
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

x2gd.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx12083m	-Xmx12083m
mapreduce.java.opts	-Xmx24166m	-Xmx24166m
mapreduce.map.memory.mb	15104	15104
mapreduce.reduce.memory.mb	30208	30208
yarn.app.mapreduce.am.resource.mb	30208	30208

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

x2gd.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.java.opts	-Xmx24576m	-Xmx24576m
mapreduce.map.memory.mb	15360	15360
mapreduce.reduce.memory.mb	30720	30720
yarn.app.mapreduce.am.resource.mb	30720	30720
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

x2gd.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.java.opts	-Xmx24712m	-Xmx24712m
mapreduce.map.memory.mb	15445	15445
mapreduce.reduce.memory.mb	30890	30890
yarn.app.mapreduce.am.resource.mb	30890	30890
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30906
yarn.nodemanager.resource.memory-mb	741376	30906

x2gd.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.java.opts	-Xmx24780m	-Xmx24780m
mapreduce.map.memory.mb	1548	1548
mapreduce.reduce.memory.mb	30976	30976
yarn.app.mapreduce.am.resource.mb	30976	30976

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

instâncias x2idn

x2idn.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.java.opts	-Xmx24780m	-Xmx24780m
mapreduce.map.memory.mb	1548	1548
mapreduce.reduce.memory.mb	30976	30976
yarn.app.mapreduce.am.resource.mb	30976	30976
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

x2idn.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx12425m	-Xmx12425m
mapreduce.java.opts	-Xmx24850m	-Xmx24850m
mapreduce.map.memory.mb	15531	15531
mapreduce.reduce.memory.mb	31062	31062
yarn.app.mapreduce.am.resource.mb	31062	31062
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31030
yarn.nodemanager.resource.memory-mb	1490944	31030

x2idn.32xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx12442m	-Xmx12442m
mapreduce.java.opts	-Xmx24884m	-Xmx24884m
mapreduce.map.memory.mb	1552	1552
mapreduce.reduce.memory.mb	31104	31104
yarn.app.mapreduce.am.resource.mb	31104	31104

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1990-656	31104
yarn.nodemanager.resource.memory-mb	1990-656	31104

instâncias x2iedn

x2iedn.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx23347m	-Xmx23347m
mapreduce.java.opts	-Xmx46694m	-Xmx46694m
mapreduce.map.memory.mb	29184	29184
mapreduce.reduce.memory.mb	58368	58368
yarn.app.mapreduce.am.resource.mb	58368	58368
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	0
yarn.nodemanager.resource.memory-mb	116736	0

x2iedn.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx24166m	-Xmx24166m
mapreduce.java.opts	-Xmx48332m	-Xmx48332m
mapreduce.map.memory.mb	30208	30208
mapreduce.reduce.memory.mb	60416	60416
yarn.app.mapreduce.am.resource.mb	60416	60416
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	0
yarn.nodemanager.resource.memory-mb	241664	0

x2iedn.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx24576m	-Xmx24576m
mapreduce.java.opts	-Xmx49152m	-Xmx49152m
mapreduce.map.memory.mb	30720	30720
mapreduce.reduce.memory.mb	61440	61440
yarn.app.mapreduce.am.resource.mb	61440	61440

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	0
yarn.nodemanager.resource.memory-mb	491520	0

x2iedn.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx24781m	-Xmx24781m
mapreduce.java.opts	-Xmx49562m	-Xmx49562m
mapreduce.map.memory.mb	30976	30976
mapreduce.reduce.memory.mb	61952	61952
yarn.app.mapreduce.am.resource.mb	61952	61952
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	0
yarn.nodemanager.resource.memory-mb	991232	0

x2iedn.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx24883m	-Xmx24883m
mapreduce.java.opts	-Xmx49766m	-Xmx49766m
mapreduce.map.memory.mb	31104	31104
mapreduce.reduce.memory.mb	62208	62208
yarn.app.mapreduce.am.resource.mb	62208	62208
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1990-656	0
yarn.nodemanager.resource.memory-mb	1990-656	0

x2iedn.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx24918m	-Xmx24918m
mapreduce.java.opts	-Xmx49836m	-Xmx49836m
mapreduce.map.memory.mb	31147	31147
mapreduce.reduce.memory.mb	6294	6294
yarn.app.mapreduce.am.resource.mb	6294	6294

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	2990080	-32
yarn.nodemanager.resource.memory-mb	2990080	-32

x2iedn.32xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx24934m	-Xmx24934m
mapreduce.java.opts	-Xmx49868m	-Xmx49868m
mapreduce.map.memory.mb	31168	31168
mapreduce.reduce.memory.mb	62336	62336
yarn.app.mapreduce.am.resource.mb	62336	62336
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	3989504	0
yarn.nodemanager.resource.memory-mb	3989504	0

instâncias x8g

x8g.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.java.opts	-Xmx21708m	-Xmx21708m
mapreduce.map.memory.mb	13568	13568
mapreduce.reduce.memory.mb	27136	27136
yarn.app.mapreduce.am.resource.mb	27136	27136
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

x8g.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.java.opts	-Xmx23348m	-Xmx23348m
mapreduce.map.memory.mb	14592	14592
mapreduce.reduce.memory.mb	29184	29184
yarn.app.mapreduce.am.resource.mb	29184	29184

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

x8g.4xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx12083m	-Xmx12083m
mapreduce.java.opts	-Xmx24166m	-Xmx24166m
mapreduce.map.memory.mb	15104	15104
mapreduce.reduce.memory.mb	30208	30208
yarn.app.mapreduce.am.resource.mb	30208	30208
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

x8g.8xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.java.opts	-Xmx24576m	-Xmx24576m
mapreduce.map.memory.mb	15360	15360
mapreduce.reduce.memory.mb	30720	30720
yarn.app.mapreduce.am.resource.mb	30720	30720
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

x8g.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.java.opts	-Xmx24712m	-Xmx24712m
mapreduce.map.memory.mb	15445	15445
mapreduce.reduce.memory.mb	30890	30890
yarn.app.mapreduce.am.resource.mb	30890	30890

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30906
yarn.nodemanager.resource.memory-mb	741376	30906

x8g.16xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.java.opts	-Xmx24780m	-Xmx24780m
mapreduce.map.memory.mb	15488	15488
mapreduce.reduce.memory.mb	30976	30976
yarn.app.mapreduce.am.resource.mb	30976	30976
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

x8g.24xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx12425m	-Xmx12425m
mapreduce.java.opts	-Xmx24850m	-Xmx24850m
mapreduce.map.memory.mb	15531	15531
mapreduce.reduce.memory.mb	31062	31062
yarn.app.mapreduce.am.resource.mb	31062	31062
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31030
yarn.nodemanager.resource.memory-mb	1490944	31030

x8g.48xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-XMX12458m	-XMX12458m
mapreduce.java.opts	-XMX24916m	-XMX24916m
mapreduce.map.memory.mb	1573	1573
mapreduce.reduce.memory.mb	31146	31146
yarn.app.mapreduce.am.resource.mb	31146	31146

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	2990080	31210
yarn.nodemanager.resource.memory-mb	2990080	31210

Instâncias z1d

z1d.xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx4915m	-Xmx4915m
mapreduce.java.opts	-Xmx9830m	-Xmx9830m
mapreduce.map.memory.mb	6144	6144
mapreduce.reduce.memory.mb	12288	12288
yarn.app.mapreduce.am.resource.mb	12288	12288
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

z1d.2xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.java.opts	-Xmx11468m	-Xmx11468m
mapreduce.map.memory.mb	7168	7168
mapreduce.reduce.memory.mb	14336	14336
yarn.app.mapreduce.am.resource.mb	14336	14336
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

z1d.3xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6007m	-Xmx6007m
mapreduce.java.opts	-Xmx12014m	-Xmx12014m
mapreduce.map.memory.mb	7509	7509
mapreduce.reduce.memory.mb	15018	15018
yarn.app.mapreduce.am.resource.mb	15018	15018

Opção de configuração	Valor padrão	Com HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	90112	30040
yarn.nodemanager.resource.memory-mb	90112	30040

z1d.6xlarge

Opção de configuração	Valor padrão	Com HBase instalado
mapreduce.map.java.opts	-Xmx6281m	-Xmx6281m
mapreduce.java.opts	-Xmx12562m	-Xmx12562m
mapreduce.map.memory.mb	7851	7851
mapreduce.reduce.memory.mb	15702	15702
yarn.app.mapreduce.am.resource.mb	15702	15702
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31396
yarn.nodemanager.resource.memory-mb	188416	31396

z1d.12xlarge

Opção de configuração	Valor padrão	Com HBase instalado
`mapreduce.map.java.opts`	-Xmx6417m	-Xmx6417m
`mapreduce.java.opts`	-Xmx12834m	-Xmx12834m
`mapreduce.map.memory.mb`	8021	8021
`mapreduce.reduce.memory.mb`	16042	16042
`yarn.app.mapreduce.am.resource.mb`	16042	16042
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	385024	32100
`yarn.nodemanager.resource.memory-mb`	385024	32100

Definições de configuração de daemon do Hadoop

As configurações do daemon do Hadoop são diferentes dependendo do tipo de EC2 instância que um nó do cluster usa. As tabelas a seguir listam as configurações padrão para cada tipo de EC2 instância.

Para personalizar essas configurações, use a classificação de configuração `hadoop-env`. Para obter mais informações, consulte Configurar aplicações.

Tipos de instância

- Instâncias c1
- Instâncias c3
- Instâncias c4
- Instâncias c5

- Instâncias c5a
- Instâncias c5ad
- Instâncias c5d
- Instâncias c5n
- Instâncias c6a
- Instâncias c6g
- Instâncias c6gd
- Instâncias c6gn
- Instâncias c6i
- Instâncias c6id
- Instâncias c6in
- Instâncias c7a
- Instâncias c7g
- Instâncias c7gd
- Instâncias c7gn
- Instâncias C7i
- Instâncias c7i-flex
- Instâncias c8g
- instâncias c8gd
- Instâncias d2
- Instâncias d3
- Instâncias d3en
- instâncias f2
- Instâncias g3
- Instâncias g3s
- Instâncias do g4dn
- Instâncias g5
- Instâncias g6
- Instâncias g6e
- Instâncias gr6

- Instâncias h1
- Instâncias i2
- Instâncias i3
- Instâncias i3en
- Instâncias i4g
- Instâncias i4i
- instâncias i7i
- Instâncias i7ie
- instâncias i8g
- Instâncias im4gn
- Instâncias is4gen
- Instâncias m1
- Instâncias m2
- Instâncias m3
- Instâncias m4
- Instâncias m5
- Instâncias m5a
- Instâncias m5ad
- Instâncias m5d
- Instâncias m5dn
- Instâncias m5n
- instâncias m5zn
- Instâncias m6a
- Instâncias m6g
- Instâncias m6gd
- Instâncias m6i
- Instâncias m6id
- Instâncias m6idn
- Instâncias m6in
- Instâncias M7a

- Instâncias m7g
- Instâncias m7gd
- Instâncias m7i
- Instâncias m7i-flex
- Instâncias m8g
- instâncias m8gd
- Instâncias p2
- Instâncias p3
- Instâncias p4d
- Instâncias p5
- Instâncias r3
- Instâncias r4
- Instâncias r5
- Instâncias r5a
- Instâncias r5ad
- Instâncias r5b
- Instâncias r5d
- Instâncias r5dn
- Instâncias r5n
- Instâncias r6a
- Instâncias r6g
- Instâncias r6gd
- Instâncias r6i
- Instâncias r6iD
- instâncias r6idn
- Instâncias r6in
- Instâncias R7a
- Instâncias r7g
- Instâncias r7gd
- Instâncias r7i

- Instâncias R7iz
- Instâncias r8g
- instâncias r8gd
- Instâncias x1
- Instâncias x1e
- Instâncias x2gd
- instâncias x2idn
- instâncias x2iedn
- Instâncias X8g
- Instâncias z1d

Instâncias c1

c1.medium

Parameter	Valor
YARN_RESOURCEMANAGER_HEAPSIZE	192
YARN_PROXYSERVER_HEAPSIZE	96
YARN_NODEMANAGER_HEAPSIZE	128
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	128
HADOOP_NAMENODE_HEAPSIZE	192
HADOOP_DATANODE_HEAPSIZE	96

c1.xlarge

Parameter	Valor
YARN_RESOURCEMANAGER_HEAPSIZE	768

Parameter	Valor
YARN_PROXYSERVER_HEAPSIZE	384
YARN_NODEMANAGER_HEAPSIZE	512
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	512
HADOOP_NAMENODE_HEAPSIZE	768
HADOOP_DATANODE_HEAPSIZE	384

Instâncias c3

c3.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c3.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2396

Parameter	Valor
YARN_PROXYSERVER_HEAPSIZE	2396
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2396
HADOOP_NAMENODE_HEAPSIZE	1740
HADOOP_DATANODE_HEAPSIZE	757

c3.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2703
YARN_PROXYSERVER_HEAPSIZE	2703
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2703
HADOOP_NAMENODE_HEAPSIZE	3276
HADOOP_DATANODE_HEAPSIZE	1064

c3.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3317
YARN_PROXYSERVER_HEAPSIZE	3317

Parameter	Valor
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3317
HADOOP_NAMENODE_HEAPSIZE	6348
HADOOP_DATANODE_HEAPSIZE	1679

Instâncias c4

c4.large

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	1152
YARN_PROXYSERVER_HEAPSIZE	1152
YARN_NODEMANAGER_HEAPSIZE	1152
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	1152
HADOOP_NAMENODE_HEAPSIZE	576
HADOOP_DATANODE_HEAPSIZE	384

c4.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124

Parameter	Valor
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c4.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2396
YARN_PROXYSERVER_HEAPSIZE	2396
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2396
HADOOP_NAMENODE_HEAPSIZE	1740
HADOOP_DATANODE_HEAPSIZE	757

c4.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2703
YARN_PROXYSERVER_HEAPSIZE	2703
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2703
HADOOP_NAMENODE_HEAPSIZE	3276
HADOOP_DATANODE_HEAPSIZE	1064

c4.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3317
YARN_PROXYSERVER_HEAPSIZE	3317
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3317
HADOOP_NAMENODE_HEAPSIZE	6348
HADOOP_DATANODE_HEAPSIZE	1679

Instâncias c5

c5.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2252
YARN_PROXYSERVER_HEAPSIZE	2252
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2252
HADOOP_NAMENODE_HEAPSIZE	1024
HADOOP_DATANODE_HEAPSIZE	614

c5.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

c5.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

c5.9xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3563
YARN_PROXYSERVER_HEAPSIZE	3563
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3563
HADOOP_NAMENODE_HEAPSIZE	7577
HADOOP_DATANODE_HEAPSIZE	1925

c5.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

c5.18xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5038
YARN_PROXYSERVER_HEAPSIZE	5038
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5038
HADOOP_NAMENODE_HEAPSIZE	14950
HADOOP_DATANODE_HEAPSIZE	3399

c5.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias c5a

c5a.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c5a.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c5a.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c5a.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c5a.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

c5a.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c5a.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias c5ad

c5ad.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c5ad.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c5ad.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c5ad.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c5ad.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c5ad.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c5ad.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias c5d

c5d.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2252
YARN_PROXYSERVER_HEAPSIZE	2252
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2252
HADOOP_NAMENODE_HEAPSIZE	1024
HADOOP_DATANODE_HEAPSIZE	614

c5d.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

c5d.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

c5d.9xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3563
YARN_PROXYSERVER_HEAPSIZE	3563
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3563
HADOOP_NAMENODE_HEAPSIZE	7577
HADOOP_DATANODE_HEAPSIZE	1925

c5d.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

c5d.18xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5038
YARN_PROXYSERVER_HEAPSIZE	5038
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5038
HADOOP_NAMENODE_HEAPSIZE	14950
HADOOP_DATANODE_HEAPSIZE	3399

c5d.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias c5n

c5n.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2304
YARN_PROXYSERVER_HEAPSIZE	2304
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2304
HADOOP_NAMENODE_HEAPSIZE	1280
HADOOP_DATANODE_HEAPSIZE	665

c5n.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2519
YARN_PROXYSERVER_HEAPSIZE	2519
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2519
HADOOP_NAMENODE_HEAPSIZE	2355
HADOOP_DATANODE_HEAPSIZE	880

c5n.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2949
YARN_PROXYSERVER_HEAPSIZE	2949
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2949
HADOOP_NAMENODE_HEAPSIZE	4505
HADOOP_DATANODE_HEAPSIZE	1310

c5n.9xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

c5n.18xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias c6a

c6a.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6a.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6a.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6a.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6a.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6a.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c6a.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c6a.32xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

c6a.48xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias c6g

c6g.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6g.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6g.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6g.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6g.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6g.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

Instâncias c6gd

c6gd.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6gd.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6gd.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6gd.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6gd.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6gd.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

Instâncias c6gn

c6gn.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6gn.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6gn.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6gn.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6gn.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6gn.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

Instâncias c6i

c6i.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6i.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6i.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6i.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6i.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6i.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c6i.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c6i.32xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias c6id

c6id.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6id.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6id.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6id.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6id.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6id.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c6id.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c6id.32xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias c6in

c6in.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6in.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6in.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6in.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6in.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6in.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c6in.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c6in.32xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias c7a

c7a.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c7a.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c7a.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c7a.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c7a.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c7a.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c7a.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c7a.32xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

c7a.48xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias c7g

c7g.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c7g.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c7g.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c7g.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c7g.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c7g.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

Instâncias c7gd

c7gd.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c7gd.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c7gd.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c7gd.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c7gd.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c7gd.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

Instâncias c7gn

c7gn.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c7gn.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c7gn.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c7gn.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c7gn.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c7gn.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

Instâncias C7i

c7i.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c7i.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c7i.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c7i.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c7i.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c7i.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c7i.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c7i.48xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias c7i-flex

c7i-flex.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c7i-flex.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c7i-flex.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c7i-flex.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c7i-flex.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c7i-flex.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

Instâncias c8g

c8g.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c8g.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c8g.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c8g.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c8g.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c8g.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c8g.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c8g.48xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

instâncias c8gd

c8gd.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c8gd.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c8gd.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c8gd.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c8gd.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c8gd.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c8gd.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c8gd.48xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias d2

d2.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

d2.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

d2.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

d2.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias d3

d3.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

d3.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

d3.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

d3.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias d3en

d3en.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

d3en.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

d3en.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

d3en.6xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

d3en.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

d3en.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

instâncias f2

f2.6xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

f2.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

f2.48xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	42065
YARN_PROXYSERVER_HEAPSIZE	42065
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	42065
HADOOP_NAMENODE_HEAPSIZE	200089
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias g3

g3.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

g3.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

g3.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias g3s

g3s.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

Instâncias do g4dn

g4dn.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416

Parameter	Valor
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

g4dn.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

g4dn.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

g4dn.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

g4dn.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

g4dn.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias g5

g5.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

g5.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

g5.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

g5.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

g5.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

g5.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

g5.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

g5.48xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias g6

g6.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2396
YARN_PROXYSERVER_HEAPSIZE	2396
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2396
HADOOP_NAMENODE_HEAPSIZE	1740
HADOOP_DATANODE_HEAPSIZE	757

g6.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2703
YARN_PROXYSERVER_HEAPSIZE	2703
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2703
HADOOP_NAMENODE_HEAPSIZE	3276
HADOOP_DATANODE_HEAPSIZE	1064

g6.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3317
YARN_PROXYSERVER_HEAPSIZE	3317
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3317
HADOOP_NAMENODE_HEAPSIZE	6348
HADOOP_DATANODE_HEAPSIZE	1679

g6.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4567
YARN_PROXYSERVER_HEAPSIZE	4567
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4567
HADOOP_NAMENODE_HEAPSIZE	12595
HADOOP_DATANODE_HEAPSIZE	2928

g6.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5795
YARN_PROXYSERVER_HEAPSIZE	5795
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5795
HADOOP_NAMENODE_HEAPSIZE	18739
HADOOP_DATANODE_HEAPSIZE	4096

g6.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7045
YARN_PROXYSERVER_HEAPSIZE	7045
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7045
HADOOP_NAMENODE_HEAPSIZE	24985
HADOOP_DATANODE_HEAPSIZE	4096

g6.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9543
YARN_PROXYSERVER_HEAPSIZE	9543
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9543
HADOOP_NAMENODE_HEAPSIZE	37478
HADOOP_DATANODE_HEAPSIZE	4096

g6.48xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	16998
YARN_PROXYSERVER_HEAPSIZE	16998
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	16998
HADOOP_NAMENODE_HEAPSIZE	74752
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias g6e

g6e.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

g6e.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

g6e.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

g6e.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

g6e.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

g6e.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

g6e.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

g6e.48xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias gr6

gr6.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4567
YARN_PROXYSERVER_HEAPSIZE	4567
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4567
HADOOP_NAMENODE_HEAPSIZE	12595
HADOOP_DATANODE_HEAPSIZE	2928

gr6.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7045
YARN_PROXYSERVER_HEAPSIZE	7045
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7045
HADOOP_NAMENODE_HEAPSIZE	24985
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias h1

h1.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

h1.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

h1.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

h1.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias i2

i2.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i2.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i2.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

i2.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias i3

i3.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i3.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i3.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

i3.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

i3.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias i3en

i3en.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

i3en.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

i3en.3xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

i3en.6xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

i3en.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

i3en.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17817
YARN_PROXYSERVER_HEAPSIZE	17817
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17817
HADOOP_NAMENODE_HEAPSIZE	78848
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias i4g

i4g.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i4g.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i4g.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

i4g.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

i4g.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias i4i

i4i.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i4i.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i4i.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

i4i.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

i4i.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

i4i.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

i4i.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

i4i.32xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

instâncias i7i

i7i.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i7i.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i7i.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

i7i.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

i7i.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

i7i.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

i7i.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

i7i.48xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias i7ie

i7ie.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i7ie.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i7ie.3xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

i7ie.6xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

i7ie.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

i7ie.18xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	13332
YARN_PROXYSERVER_HEAPSIZE	13332
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	13332
HADOOP_NAMENODE_HEAPSIZE	56422
HADOOP_DATANODE_HEAPSIZE	4096

i7ie.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

i7ie.48xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

instâncias i8g

i8g.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i8g.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i8g.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

i8g.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

i8g.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

i8g.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

i8g.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

i8g.48xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias im4gn

im4gn.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

im4gn.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

im4gn.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

im4gn.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

im4gn.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias is4gen

is4gen.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2557
YARN_PROXYSERVER_HEAPSIZE	2557
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2557
HADOOP_NAMENODE_HEAPSIZE	2547
HADOOP_DATANODE_HEAPSIZE	919

is4gen.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3025
YARN_PROXYSERVER_HEAPSIZE	3025
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3025
HADOOP_NAMENODE_HEAPSIZE	4889
HADOOP_DATANODE_HEAPSIZE	1387

is4gen.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

is4gen.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias m1

m1.small

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	256
YARN_PROXYSERVER_HEAPSIZE	96
YARN_NODEMANAGER_HEAPSIZE	192
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	128
HADOOP_NAMENODE_HEAPSIZE	192
HADOOP_DATANODE_HEAPSIZE	96

m1.medium

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	384
YARN_PROXYSERVER_HEAPSIZE	192
YARN_NODEMANAGER_HEAPSIZE	256

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	256
HADOOP_NAMENODE_HEAPSIZE	384
HADOOP_DATANODE_HEAPSIZE	192

m1.large

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	768
YARN_PROXYSERVER_HEAPSIZE	384
YARN_NODEMANAGER_HEAPSIZE	512
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	512
HADOOP_NAMENODE_HEAPSIZE	768
HADOOP_DATANODE_HEAPSIZE	384

m1.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	1024
YARN_PROXYSERVER_HEAPSIZE	512
YARN_NODEMANAGER_HEAPSIZE	768

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	1024
HADOOP_NAMENODE_HEAPSIZE	2304
HADOOP_DATANODE_HEAPSIZE	384

Instâncias m2

m2.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	1536
YARN_PROXYSERVER_HEAPSIZE	1024
YARN_NODEMANAGER_HEAPSIZE	1024
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	1024
HADOOP_NAMENODE_HEAPSIZE	3072
HADOOP_DATANODE_HEAPSIZE	384

m2.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	1536
YARN_PROXYSERVER_HEAPSIZE	1024
YARN_NODEMANAGER_HEAPSIZE	1024

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	1536
HADOOP_NAMENODE_HEAPSIZE	6144
HADOOP_DATANODE_HEAPSIZE	384

m2.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2048
YARN_PROXYSERVER_HEAPSIZE	1024
YARN_NODEMANAGER_HEAPSIZE	1536
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	1536
HADOOP_NAMENODE_HEAPSIZE	12288
HADOOP_DATANODE_HEAPSIZE	384

Instâncias m3

m3.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2396
YARN_PROXYSERVER_HEAPSIZE	2396
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2396
HADOOP_NAMENODE_HEAPSIZE	1740
HADOOP_DATANODE_HEAPSIZE	757

m3.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2703
YARN_PROXYSERVER_HEAPSIZE	2703
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2703
HADOOP_NAMENODE_HEAPSIZE	3276
HADOOP_DATANODE_HEAPSIZE	1064

Instâncias m4

m4.large

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2252
YARN_PROXYSERVER_HEAPSIZE	2252
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2252
HADOOP_NAMENODE_HEAPSIZE	1024
HADOOP_DATANODE_HEAPSIZE	614

m4.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

m4.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

m4.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

m4.10xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5365
YARN_PROXYSERVER_HEAPSIZE	5365
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5365
HADOOP_NAMENODE_HEAPSIZE	16588
HADOOP_DATANODE_HEAPSIZE	3727

m4.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias m5

m5.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

m5.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

m5.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

m5.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

m5.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

m5.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

m5.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias m5a

m5a.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

m5a.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

m5a.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

m5a.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

m5a.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

m5a.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

m5a.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias m5ad

m5ad.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m5ad.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m5ad.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m5ad.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m5ad.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m5ad.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m5ad.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias m5d

m5d.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

m5d.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

m5d.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

m5d.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

m5d.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

m5d.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

m5d.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias m5dn

m5dn.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m5dn.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m5dn.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m5dn.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m5dn.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m5dn.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m5dn.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias m5n

m5n.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m5n.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m5n.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m5n.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m5n.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m5n.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m5n.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

instâncias m5zn

m5zn.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2396
YARN_PROXYSERVER_HEAPSIZE	2396
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2396
HADOOP_NAMENODE_HEAPSIZE	1740
HADOOP_DATANODE_HEAPSIZE	757

m5zn.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m5zn.3xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3025
YARN_PROXYSERVER_HEAPSIZE	3025
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3025
HADOOP_NAMENODE_HEAPSIZE	4889
HADOOP_DATANODE_HEAPSIZE	1387

m5zn.6xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

m5zn.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias m6a

m6a.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6a.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6a.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6a.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6a.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m6a.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m6a.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m6a.32xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

m6a.48xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias m6g

m6g.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6g.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6g.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6g.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6g.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

m6g.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias m6gd

m6gd.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6gd.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6gd.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6gd.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6gd.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

m6gd.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias m6i

m6i.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6i.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6i.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6i.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6i.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

m6i.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m6i.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m6i.32xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias m6id

m6id.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6id.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6id.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6id.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6id.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m6id.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m6id.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m6id.32xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias m6idn

m6idn.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6idn.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6idn.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6idn.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6idn.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m6idn.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m6idn.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m6idn.32xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias m6in

m6in.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6in.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6in.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6in.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6in.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

m6in.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m6in.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m6in.32xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias M7a

m7a.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m7a.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m7a.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m7a.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m7a.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m7a.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m7a.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m7a.32xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

m7a.48xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias m7g

m7g.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m7g.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m7g.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m7g.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m7g.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m7g.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias m7gd

m7gd.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m7gd.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m7gd.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m7gd.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m7gd.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m7gd.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias m7i

m7i.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m7i.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m7i.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m7i.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m7i.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

m7i.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m7i.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m7i.48xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias m7i-flex

m7i-flex.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m7i-flex.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m7i-flex.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m7i-flex.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m7i-flex.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m7i-flex.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias m8g

m8g.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m8g.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m8g.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m8g.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m8g.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

m8g.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m8g.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m8g.48xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

instâncias m8gd

m8gd.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m8gd.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m8gd.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m8gd.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m8gd.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m8gd.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m8gd.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m8gd.48xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias p2

p2.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

p2.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

p2.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias p3

p3.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

p3.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

p3.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias p4d

p4d.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	24576
YARN_PROXYSERVER_HEAPSIZE	24576
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	24576
HADOOP_NAMENODE_HEAPSIZE	112640
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias p5

p5.48xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	42065
YARN_PROXYSERVER_HEAPSIZE	42065

Parameter	Valor
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	42065
HADOOP_NAMENODE_HEAPSIZE	200089
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias r3

r3.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r3.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338

Parameter	Valor
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r3.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r3.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias r4

r4.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r4.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r4.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r4.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r4.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias r5

r5.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

r5.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

r5.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

r5.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

r5.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

r5.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12574
YARN_PROXYSERVER_HEAPSIZE	12574
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12574
HADOOP_NAMENODE_HEAPSIZE	52633
HADOOP_DATANODE_HEAPSIZE	4096

r5.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17817
YARN_PROXYSERVER_HEAPSIZE	17817
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17817
HADOOP_NAMENODE_HEAPSIZE	78848
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias r5a

r5a.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

r5a.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

r5a.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

r5a.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

r5a.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

r5a.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12574
YARN_PROXYSERVER_HEAPSIZE	12574
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12574
HADOOP_NAMENODE_HEAPSIZE	52633
HADOOP_DATANODE_HEAPSIZE	4096

r5a.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17817
YARN_PROXYSERVER_HEAPSIZE	17817
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17817
HADOOP_NAMENODE_HEAPSIZE	78848
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias r5ad

r5ad.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r5ad.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r5ad.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r5ad.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r5ad.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r5ad.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12247
YARN_PROXYSERVER_HEAPSIZE	12247
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12247
HADOOP_NAMENODE_HEAPSIZE	50995
HADOOP_DATANODE_HEAPSIZE	4096

r5ad.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias r5b

r5b.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r5n.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r5b.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r5b.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r5b.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r5b.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r5b.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias r5d

r5d.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

r5d.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

r5d.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

r5d.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

r5d.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

r5d.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12574
YARN_PROXYSERVER_HEAPSIZE	12574
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12574
HADOOP_NAMENODE_HEAPSIZE	52633
HADOOP_DATANODE_HEAPSIZE	4096

r5d.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17817
YARN_PROXYSERVER_HEAPSIZE	17817
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17817
HADOOP_NAMENODE_HEAPSIZE	78848
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias r5dn

r5dn.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r5dn.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r5dn.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r5dn.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r5dn.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r5dn.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r5dn.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias r5n

r5n.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r5n.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r5n.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r5n.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r5n.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r5n.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r5n.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias r6a

r6a.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6a.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6a.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6a.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6a.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6a.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r6a.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r6a.32xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

r6a.48xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias r6g

r6g.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6g.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6g.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6g.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6g.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6g.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias r6gd

r6gd.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6gd.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6gd.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6gd.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6gd.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6gd.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias r6i

r6i.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6i.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6i.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6i.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6i.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6i.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r6i.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r6i.32xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	21544
YARN_PROXYSERVER_HEAPSIZE	21544
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	21544
HADOOP_NAMENODE_HEAPSIZE	97484
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias r6iD

r6id.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6id.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6id.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6id.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6id.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6id.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r6id.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r6id.32xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

instâncias r6idn

r6idn.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6idn.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6idn.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6idn.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6idn.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6idn.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r6idn.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r6idn.32xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias r6in

r6in.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6in.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6in.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6in.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6in.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6in.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r6in.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r6in.32xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias R7a

r7a.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r7a.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r7a.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r7a.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r7a.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r7a.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r7a.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r7a.32xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

r7a.48xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias r7g

r7g.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r7g.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r7g.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r7g.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r7g.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r7g.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias r7gd

r7gd.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r7gd.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r7gd.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r7gd.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r7gd.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r7gd.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias r7i

r7i.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r7i.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r7i.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r7i.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r7i.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r7i.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r7i.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r7i.48xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias R7iz

r7iz.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r7iz.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r7iz.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r7iz.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r7iz.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r7iz.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r7iz.32xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	21544
YARN_PROXYSERVER_HEAPSIZE	21544
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	21544
HADOOP_NAMENODE_HEAPSIZE	97484
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias r8g

r8g.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r8g.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r8g.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r8g.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r8g.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r8g.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r8g.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r8g.48xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

instâncias r8gd

r8gd.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r8gd.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r8gd.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r8gd.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r8gd.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r8gd.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r8gd.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r8gd.48xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias x1

x1.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	21544
YARN_PROXYSERVER_HEAPSIZE	21544
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	21544
HADOOP_NAMENODE_HEAPSIZE	97484
HADOOP_DATANODE_HEAPSIZE	4096

x1.32xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	41000
YARN_PROXYSERVER_HEAPSIZE	41000
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	41000
HADOOP_NAMENODE_HEAPSIZE	194764
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias x1e

x1e.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4520
YARN_PROXYSERVER_HEAPSIZE	4520
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4520
HADOOP_NAMENODE_HEAPSIZE	12364
HADOOP_DATANODE_HEAPSIZE	2882

x1e.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	6952
YARN_PROXYSERVER_HEAPSIZE	6952
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6952
HADOOP_NAMENODE_HEAPSIZE	24524
HADOOP_DATANODE_HEAPSIZE	4096

x1e.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	11816
YARN_PROXYSERVER_HEAPSIZE	11816
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	11816
HADOOP_NAMENODE_HEAPSIZE	48844
HADOOP_DATANODE_HEAPSIZE	4096

x1e.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	21544
YARN_PROXYSERVER_HEAPSIZE	21544
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	21544
HADOOP_NAMENODE_HEAPSIZE	97484
HADOOP_DATANODE_HEAPSIZE	4096

x1e.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	41000
YARN_PROXYSERVER_HEAPSIZE	41000
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	41000
HADOOP_NAMENODE_HEAPSIZE	194764
HADOOP_DATANODE_HEAPSIZE	4096

x1e.32xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	79912
YARN_PROXYSERVER_HEAPSIZE	79912
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	79912
HADOOP_NAMENODE_HEAPSIZE	389324
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias x2gd

x2gd.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

x2gd.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

x2gd.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

x2gd.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

x2gd.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

x2gd.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

instâncias x2idn

x2idn.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

x2idn.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

x2idn.32xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	42065
YARN_PROXYSERVER_HEAPSIZE	42065
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	42065
HADOOP_NAMENODE_HEAPSIZE	200089
HADOOP_DATANODE_HEAPSIZE	4096

instâncias x2iedn

x2iedn.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

x2iedn.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

x2iedn.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

x2iedn.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

x2iedn.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	42065
YARN_PROXYSERVER_HEAPSIZE	42065
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	42065
HADOOP_NAMENODE_HEAPSIZE	200089
HADOOP_DATANODE_HEAPSIZE	4096

x2iedn.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	62054
YARN_PROXYSERVER_HEAPSIZE	62054
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	62054
HADOOP_NAMENODE_HEAPSIZE	300032
HADOOP_DATANODE_HEAPSIZE	4096

x2iedn.32xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	82042
YARN_PROXYSERVER_HEAPSIZE	82042
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	82042
HADOOP_NAMENODE_HEAPSIZE	399974
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias X8g

x8g.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

x8g.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

x8g.4xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

x8g.8xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

x8g.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

x8g.16xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

x8g.24xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

x8g.48xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	62054
YARN_PROXYSERVER_HEAPSIZE	62054
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	62054
HADOOP_NAMENODE_HEAPSIZE	300032
HADOOP_DATANODE_HEAPSIZE	4096

Instâncias z1d

z1d.xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

z1d.2xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

z1d.3xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

z1d.6xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

z1d.12xlarge

Parameter	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048

Parameter	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

Configuração do HDFS

A tabela a seguir descreve os parâmetros padrão do Hadoop Distributed File System (HDFS) e suas configurações. Você pode alterar esses valores usando a classificação de configuração `hdfs-site`. Para obter mais informações, consulte Configurar aplicações.

Warning

1. Definir `dfs.replication` como 1 em clusters com menos de quatro nós poderá causar perda de dados do HDFS se um único nó ficar inativo. Se seu cluster tiver armazenamento HDFS, é recomendável configurar o cluster com pelo menos quatro nós principais para workloads de produção para evitar perda de dados.
2. O Amazon EMR não permitirá que os clusters escalem os nós principais abaixo de `dfs.replication`. Por exemplo, se `dfs.replication = 2`, o número mínimo de nós central será 2.
3. Ao usar o Ajuste de Escala Gerenciado, o ajuste de escala automático ou optar por redimensionar manualmente o cluster, é recomendável definir `dfs.replication` como 2 ou mais.

Parameter	Definição	Valor padrão
`dfs.block.size`	O tamanho de blocos HDFS. Ao trabalhar com dados armazenados no HDFS, o tamanho da divisão, geralmente, é o tamanho de um bloco HDFS. Números maiores fornecem menos granularidade de tarefas, mas também impõem menos pressão no cluster `NameNode`.	134217728 (128 MB)

Parameter	Definição	Valor padrão
`dfs.replication`	O número de cópias de cada bloco a ser armazenado para durabilidade. O Amazon EMR define esse valor com base no número de nós centrais com os quais o cluster é provisionado. Ajuste o valor para atender a suas necessidades. Para substituir o valor padrão, use a classificação `hdfs-site` .	1 para clusters provisionados com menos de quatro nós centrais 2 para clusters provisionados com menos de dez nós centrais 3 para todos os outros clusters

Criptografia transparente no HDFS com o Amazon EMR

A criptografia transparente é implementada com o uso de zonas de criptografia HDFS, que são caminhos HDFS que você mesmo define. Cada zona de criptografia tem sua própria chave, que é armazenada no servidor de chaves especificado usando a classificação da configuração `hdfs-site`.

A partir do Amazon EMR versão 4.8.0, você pode usar as configurações de segurança do Amazon EMR para definir configurações de criptografia de dados para clusters com mais facilidade. Configurações de segurança oferecem configurações para habilitar a segurança dos dados em trânsito e dos dados em repouso em volumes de armazenamento do Amazon Elastic Block Store (Amazon EBS) e dados do EMRFS no Amazon S3. Para obter mais informações, consulte Encrypt data in transit and at rest, no Guia de gerenciamento do Amazon EMR.

O Amazon EMR usa o Hadoop KMS por padrão; no entanto, você pode usar outro KMS que implemente a operação da API. KeyProvider Cada arquivo em uma zona de criptografia HDFS tem sua própria chave de criptografia de dados exclusiva, que é criptografada pela chave de zona de criptografia. Os dados do HDFS são criptografados end-to-end (em repouso e em trânsito) quando são gravados em uma zona de criptografia, pois as atividades de criptografia e descriptografia ocorrem somente no cliente.

Você não pode mover arquivos entre zonas de criptografias ou de uma zona de criptografia para caminhos não criptografados.

O cliente NameNode e o HDFS interagem com o Hadoop KMS (ou um KMS alternativo que você configurou) por meio da operação da API. KeyProvider O KMS é responsável por armazenar chaves de criptografia no repositório de chaves de suporte. Além disso, o Amazon EMR inclui a política de força ilimitada JCE e, portanto, você pode criar chaves com o comprimento desejado.

Para obter mais informações, consulte Transparent encryption in HDFS na documentação do Hadoop.

> **Note**
>
> No Amazon EMR, o KMS via HTTPS não está habilitado por padrão com o Hadoop KMS. Para obter mais informações sobre como habilitar o KMS por HTTPS, consulte a Documentação do Hadoop KMS.

Configurar a criptografia transparente do HDFS

É possível configurar a criptografia transparente no Amazon EMR criando chaves e adicionando zonas de criptografia. Você pode fazer isso de várias maneiras:

- Usar a operação da API de configuração do Amazon EMR ao criar um cluster
- Usando uma etapa Hadoop JAR com command-runner.jar
- Fazendo login no nó principal do cluster Hadoop e usando os clientes de linha de comando `hadoop key` e `hdfs crypto`
- Usando o REST APIs para Hadoop KMS e HDFS

Para obter mais informações sobre o REST APIs, consulte a respectiva documentação do Hadoop KMS e do HDFS.

Para criar zonas de criptografia e suas chaves na criação de cluster usando a CLI

A classificação`hdfs-encryption-zones` na operação da API de configuração permite que você especifique um nome de chave e uma zona de criptografia ao criar um cluster. O Amazon EMR cria essa chave no Hadoop KMS no cluster e configura a zona de criptografia.

- Crie um cluster com o comando a seguir.

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge --
instance-count 2 \
--applications Name=App1 Name=App2 --configurations https://s3.amazonaws.com/amzn-
s3-demo-bucket/myfolder/myConfig.json
```

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

`myConfig.json`:

```
[
 {
  "Classification": "hdfs-encryption-zones",
  "Properties": {
   "/myHDFSPath1": "path1_key",
   "/myHDFSPath2": "path2_key"
  }
 }
]
```

Para criar zonas de criptografia e suas chaves manualmente no nó principal

1. Inicie o cluster usando uma versão do Amazon EMR superior a 4.1.0.
2. Conecte-se ao nó principal do cluster com SSH.
3. Crie uma chave no Hadoop KMS.

```
$ hadoop key create path2_key
path2_key has been successfully created with options Options{cipher='AES/CTR/
NoPadding', bitLength=256, description='null', attributes=null}.
KMSClientProvider[http://ip-x-x-x-x.ec2.internal:16000/kms/v1/] has been updated.
```

> **Important**
>
> O Hadoop KMS exige que os nomes das suas chaves estejam em minúsculas. Se você usar uma chave que tenha caracteres em maiúsculas, seu cluster falhará durante a execução.

4. Crie o caminho da zona de criptografia no HDFS.

```
$ hadoop fs -mkdir /myHDFSPath2
```

5. Transforme o caminho do HDFS em uma zona de criptografia usando a chave que você criou.

```
$ hdfs crypto -createZone -keyName path2_key -path /myHDFSPath2
Added encryption zone /myHDFSPath2
```

Para criar zonas de criptografia e suas chaves manualmente usando o AWS CLI

- Adicione etapas para criar as chaves do KMS e as zonas de criptografia manualmente com o comando a seguir.

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF --steps Type=CUSTOM_JAR,Name="Create
 First Hadoop KMS Key",Jar="command-runner.jar",ActionOnFailure=CONTINUE,Args=[/
bin/bash,-c,"\"hadoop key create path1_key\""] \
Type=CUSTOM_JAR,Name="Create First Hadoop HDFS Path",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[/bin/bash,-c,"\"hadoop fs -mkdir /
myHDFSPath1\""] \
Type=CUSTOM_JAR,Name="Create First Encryption Zone",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[/bin/bash,-c,"\"hdfs crypto -createZone
 -keyName path1_key -path /myHDFSPath1\""] \
Type=CUSTOM_JAR,Name="Create Second Hadoop KMS Key",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[/bin/bash,-c,"\"hadoop key create
 path2_key\""] \
Type=CUSTOM_JAR,Name="Create Second Hadoop HDFS Path",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[/bin/bash,-c,"\"hadoop fs -mkdir /
myHDFSPath2\""] \
Type=CUSTOM_JAR,Name="Create Second Encryption Zone",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[/bin/bash,-c,"\"hdfs crypto -createZone
 -keyName path2_key -path /myHDFSPath2\""]
```

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

Considerações para criptografia transparente do HDFS

Uma prática recomendada é criar uma zona de criptografia para cada aplicativo, na qual eles podem gravar arquivos. Além disso, você pode criptografar todo o HDFS usando a hdfs-encryption-zones classificação na API de configuração e especificar o caminho raiz (/) como a zona de criptografia.

Servidor de gerenciamento de chaves do Hadoop

O Hadoop KMS é um servidor de gerenciamento de chaves que fornece a capacidade de implementar serviços criptográficos para clusters Hadoop e pode servir como fornecedor de caves para o Criptografia transparente no HDFS com o Amazon EMR. O Hadoop KMS no Amazon EMR é instalado e habilitado por padrão quando você seleciona a aplicação Hadoop ao executar um cluster do EMR. O Hadoop KMS não armazena as chaves por si só, exceto no caso de um armazenamento em cache temporário. O KMS do Hadoop atua como proxy entre o provedor de chaves e o agente de confiança do cliente para um repositório de chaves de suporte. Ele não é um repositório de chaves. O repositório de chaves padrão criado para o Hadoop KMS é a Java Cryptography Extension (JCEKS). KeyStore A política de força ilimitada JCE também está incluída e, portanto, é possível criar chaves com o comprimento desejado. O Hadoop KMS também oferece suporte a uma variedade de dispositivos ACLs que controlam o acesso às chaves e às operações de chaves, independentemente de outros aplicativos cliente, como o HDFS. O comprimento de chave padrão do Amazon EMR é 256 bits.

Para configurar o Hadoop KMS, use a hadoop-kms-site classificação para alterar as configurações. Para configurar ACLs, você usa a classificação kms-acls.

Para obter mais informações, consulte a documentação do Hadoop KMS. O Hadoop KMS é usado na criptografia transparente do Hadoop HDFS. Para saber mais sobre a criptografia transparente do HDFS, consulte o tópico HDFS transparent encryption na documentação do Apache Hadoop.

> **Note**
>
> No Amazon EMR, o KMS via HTTPS não está habilitado por padrão com o Hadoop KMS. Para saber como habilitar o KMS por HTTPS, consulte a Documentação do Hadoop KMS.

> **Important**
>
> O Hadoop KMS exige que os nomes das suas chaves estejam em minúsculas. Se você usar uma chave que tenha caracteres em maiúsculas, seu cluster falhará durante a execução.

Configurar o Hadoop KMS no Amazon EMR

No Amazon EMR versão 4.6.0 ou posteriores, `kms-http-port` é 9700 e `kms-admin-port` é 9701.

Você pode configurar o Hadoop KMS durante a criação do cluster usando a API de configuração para versões do Amazon EMR. Veja a seguir as classificações de objeto de configuração disponíveis para o Hadoop KMS:

Classificações de configuração do KMS do Hadoop

Classificação	Nome do arquivo
hadoop-kms-site	`kms-site.xml`
hadoop-kms-acls	`kms-acls.xml`
hadoop-kms-env	`kms-env.sh`
hadoop-kms-log4j	`kms-log4j.properties`

Para definir o Hadoop KMS ACLs usando a CLI

- Crie um cluster com o Hadoop KMS ACLs usando o seguinte comando:

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge --
instance-count 2 \
```

```
--applications Name=App1 Name=App2 --configurations https://s3.amazonaws.com/amzn-
s3-demo-bucket/myfolder/myConfig.json
```

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

`myConfig.json`:

```
[
 {
  "Classification": "hadoop-kms-acls",
  "Properties": {
   "hadoop.kms.blacklist.CREATE": "hdfs,foo,myBannedUser",
   "hadoop.kms.acl.ROLLOVER": "myAllowedUser"
  }
 }
]
```

Para desabilitar o cache do Hadoop KMS usando a CLI

- Crie um cluster com o Hadoop KMS `hadoop.kms.cache.enable` definido como `false`, usando o seguinte comando:

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge --
instance-count 2 \
--applications Name=App1 Name=App2 --configurations https://s3.amazonaws.com/amzn-
s3-demo-bucket/myfolder/myConfig.json
```

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

`myConfig.json`:

```
[
 {
  "Classification": "hadoop-kms-site",
  "Properties": {
   "hadoop.kms.cache.enable": "false"
  }
 }
]
```

Para definir variáveis de ambiente no script **`kms-env.sh`** usando a CLI

- Altere configurações em `kms-env.sh` por meio da configuração `hadoop-kms-env`. Crie um cluster com o Hadoop KMS usando o seguinte comando:

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge --
instance-count 2 \
--applications Name=App1 Name=App2 --configurations https://s3.amazonaws.com/amzn-
s3-demo-bucket/myfolder/myConfig.json
```

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

`myConfig.json`:

```
[
 {
  "Classification": "hadoop-kms-env",
  "Properties": {},
  "Configurations": [
   {
    "Classification": "export",
    "Properties": {
```

```
      "JAVA_LIBRARY_PATH": "/path/to/files",
      "KMS_SSL_KEYSTORE_FILE": "/non/Default/Path/.keystore",
      "KMS_SSL_KEYSTORE_PASS": "myPass"
     },
     "Configurations": []
    }
   ]
  }
]
```

Para obter informações sobre como configurar o Hadoop KMS, consulte a documentação do Hadoop KMS.

Criptografia transparente do HDFS em clusters do EMR com múltiplos nós principais

O Apache Ranger KMS é usado em um cluster do Amazon EMR com múltiplos nós centrais para criptografia transparente no HDFS.

O Apache Ranger KMS armazena a chave raiz e as chaves de zona de criptografia (EZ) em seu Amazon RDS para um cluster do Amazon EMR com múltiplos nós primários. Para habilitar a criptografia transparente no HDFS em um cluster do Amazon EMR com múltiplos nós primários, você deverá fornecer as configurações a seguir.

- O Amazon RDS ou seu próprio URL de conexão do servidor MySQL para armazenar a chave raiz do Ranger KMS e a chave da EZ
- Nome de usuário e senha do MySQL
- Senha para chave raiz do KMS do Ranger
- Arquivo PEM da autoridade de certificação (CA) para conexão SSL ao servidor MySQL. Você pode baixar o pacote de certificados para você em Baixar pacotes Região da AWS de certificados para Amazon RDS.

Você pode fornecer essas configurações usando as classificações `ranger-kms-dbks-site` e `ranger-kms-db-ca`, como demonstra o exemplo a seguir.

```
[
 {
```

```
  "Classification": "ranger-kms-dbks-site",
  "Properties": {
   "ranger.ks.jpa.jdbc.url": "jdbc:log4jdbc:mysql://mysql-host-url.xx-
xxx-1.xxx.amazonaws.com:3306/rangerkms",
   "ranger.ks.jpa.jdbc.user": "mysql-user-name",
   "ranger.ks.jpa.jdbc.password": "mysql-password",
   "ranger.db.encrypt.key.password": "password-for-encrypting-a-master-key"
  }
 },
 {
  "Classification": "ranger-kms-db-ca",
  "Properties": {
   "ranger.kms.trust.ca.file.s3.url": "<S3-path-of-downloaded-pem-file>"
  }
 }
]
```

Veja a seguir as classificações de objetos de configuração para o Apache Ranger KMS.

Classificações de configuração do KMS do Hadoop

Classificação	Descrição
ranger-kms-dbks-site	Altere os valores no arquivo dbks-site.xml do Ranger KMS.
ranger-kms-site	Altere os valores no ranger-kms-site arquivo.xml do Ranger KMS.
ranger-kms-env	Altere os valores no ambiente do Ranger KMS.
ranger-kms-log4j	Altere os valores no arquivo kms-log4j.properties do Ranger KMS.
ranger-kms-db-ca	Altere os valores do arquivo CA no S3 para conexão SSL do MySQL com o Ranger KMS.

Considerações

- É altamente recomendável criptografar sua instância do Amazon RDS para melhorar a segurança. Para obter mais informações, consulte Visão geral da criptografia de recursos do Amazon RDS.

- É altamente recomendável que você use um banco de dados MySQL separado para cada cluster do Amazon EMR com múltiplos nós primários a fim de criar barreiras de alta segurança.
- Para configurar a criptografia transparente no HDFS em um cluster do Amazon EMR com múltiplos nós primários, você deverá especificar a classificação `hdfs-encryption-zones` ao criar o cluster. Caso contrário, o Ranger KMS não será configurado ou iniciado. Não há suporte para reconfigurar a classificação `hdfs-encryption-zones` ou qualquer uma das classificações de configuração em um cluster em execução no cluster do Amazon EMR com múltiplos nós primários.
- O pacote de certificados PEM baixados em Download certificate bundles for Amazon RDS agrupa vários certificados em um arquivo. O Amazon EMR 7.3.0 e versões posteriores oferecem suporte à importação de vários certificados do arquivo PEM com a configuração `ranger.kms.trust.ca.file.s3.url`.

Criar ou executar uma aplicação Hadoop

Tópicos

- Criar binários usando o Amazon EMR
- Processamento de dados com streaming
- Processar dados com um JAR personalizado

Criar binários usando o Amazon EMR

Você pode usar o Amazon EMR como um ambiente de criação para compilar programas que serão usados no cluster. Os programas que você usa com o Amazon EMR devem ser compilados em um sistema que executa a mesma versão do Linux usada pelo Amazon EMR. Para uma versão de 32 bits, você deve compilar em uma máquina de 32 bits ou em uma máquina com as opções de compilação cruzada de 32 bits ativadas. Para uma versão de 64 bits, você deve compilar em uma máquina de 64 bits ou em uma máquina com as opções de compilação cruzada de 64 bits ativadas. Para obter mais informações sobre versões de EC2 instâncias, consulte Planejar e configurar EC2 instâncias no Guia de gerenciamento do Amazon EMR. As linguagens de programação compatíveis incluem C++, Python e C#.

A tabela a seguir descreve as etapas envolvidas na criação e no teste da sua aplicação usando o Amazon EMR.

Processo para a criação de um módulo

1	Conecte-se ao nó principal do seu cluster.
2	Copie os arquivos de código-fonte para o nó principal.
3	Crie os arquivos binários com as otimizações necessárias.
4	Copie os arquivos binários do nó principal para o Amazon S3.

Os detalhes de cada uma dessas etapas são abordados nas seções a seguir.

Para se conectar ao nó principal do cluster

- Siga as instruções em Conectar-se ao nó principal usando SSH no Guia de gerenciamento do Amazon EMR.

Para copiar os arquivos de código-fonte para o nó principal

1. Coloque os arquivos de origem em um bucket do Amazon S3. Para aprender como criar buckets e como transferir dados para o Amazon S3, consulte o Guia do usuário do Amazon Simple Storage Service.
2. Crie uma pasta no cluster do Hadoop para seus arquivos de código-fonte digitando um comando semelhante ao seguinte:

   ```
   mkdir SourceFiles
   ```
3. Copie os arquivos de origem do Amazon S3 para o nó principal digitando um comando semelhante ao seguinte:

   ```
   hadoop fs -get s3://amzn-s3-demo-bucket/SourceFiles SourceFiles
   ```

Crie os arquivos binários com as otimizações necessárias

Como você cria seus arquivos binários depende de vários fatores. Siga as instruções específicas das suas ferramentas de criação para instalar e configurar seu ambiente. Você pode usar os comandos de especificação do sistema do Hadoop para obter as informações do cluster que vão determinar como instalar o seu ambiente de criação.

Para identificar as especificações do sistema

- Use os comandos a seguir para verificar a arquitetura que você está usando para criar seus arquivos binários.

  a. Para visualizar a versão do Debian, insira o seguinte comando:

  ```
  master$ cat /etc/issue
  ```

  A saída será semelhante à seguinte.

  ```
  Debian GNU/Linux 5.0
  ```

  b. Para visualizar o nome DNS público e o tamanho do processador, insira o seguinte comando:

  ```
  master$ uname -a
  ```

  A saída será semelhante à seguinte.

  ```
  Linux domU-12-31-39-17-29-39.compute-1.internal 2.6.21.7-2.fc8xen #1 SMP Fri
   Feb 15 12:34:28 EST 2008 x86_64 GNU/Linux
  ```

  c. Para visualizar a velocidade do processador, insira o seguinte comando:

  ```
  master$ cat /proc/cpuinfo
  ```

  A saída será semelhante à seguinte.

  ```
  processor : 0
  vendor_id : GenuineIntel
  model name : Intel(R) Xeon(R) CPU E5430 @ 2.66GHz
  flags : fpu tsc msr pae mce cx8 apic mca cmov pat pse36 clflush dts acpi mmx
   fxsr sse sse2 ss ht tm syscall nx lm constant_tsc pni monitor ds_cpl vmx est
   tm2 ssse3 cx16 xtpr cda lahf_lm
  ...
  ```

Após a criação dos binários, você pode copiar os arquivos para o Amazon S3.

Copiar os binários do nó principal para o Amazon S3

- Digite o comando a seguir para copiar os binários para o bucket do Amazon S3:

```
hadoop fs -put BinaryFiles s3://amzn-s3-demo-bucket/BinaryDestination
```

Processamento de dados com streaming

O Hadoop Streaming é um utilitário que acompanha o Hadoop que permite o desenvolvimento de MapReduce executáveis em linguagens diferentes de Java. O Streaming é implementado como um arquivo JAR para que você possa executá-lo a na API do Amazon EMR ou na linha de comando, da mesma forma que executa um arquivo JAR padrão.

Esta seção descreve como usar streaming com o Amazon EMR.

> **Note**
>
> O Apache Hadoop Streaming é uma ferramenta independente. Assim sendo, não descreveremos suas funções e parâmetros aqui. Para obter mais informações sobre o Hadoop Streaming, acesse http://hadoop.apache. org/docs/stable/hadoop-streaming/HadoopStreaming.html.

Utilização do utilitário Hadoop Streaming

Esta seção descreve como usar o utilitário Hadoop Streaming.

Processo do Hadoop

1	Escreva os programas executáveis do mapeador e do reducer na linguagem de programação de sua escolha. Siga as instruções na documentação do Hadoop para escrever seus executáveis de streaming. Os programas devem ler os dados de entrada de uma entrada padrão e escrever os dados de saída em uma saída padrão. Por padrão, cada linha de entrada/saída representa um registro e o primeiro caractere de tabulação em cada linha é usado como um separador entre a chave e o valor.
2	Teste os executáveis localmente e carregue-os para o Amazon S3.

3	Use a interface da linha de comando do Amazon EMR ou o console do Amazon EMR para executar a aplicação.

Um script de mapeador é executado como um processo separado no cluster. Um executável do reducer transforma a saída de um executável do mapeador em uma saída de dados do fluxo de trabalho.

Os parâmetros `input`, `output`, `mapper` e `reducer` são necessários na maioria das aplicações de streaming. A tabela a seguir descreve esses e outros parâmetros opcionais.

Parameter	Descrição	Obrigatório
-input	Local dos dados de entrada no Amazon S3. Tipo: string Padrão: nenhum Restrição: URI. Se não há protocolo especificado, ele usa o sistema de arquivos padrão do cluster.	Sim
-output	Local onde o Amazon EMR carrega os dados processados no Amazon S3. Tipo: string Padrão: nenhum Restrição: URI Padrão: se não há um local especificado, o Amazon EMR carrega os dados no local especificado pela `input`.	Sim
-mapper	Nome do executável do mapeador. Tipo: string	Sim

Parameter	Descrição	Obrigatório
	Padrão: nenhum	
-reducer	Nome do executável do reducer. Tipo: string Padrão: nenhum	Sim
-cacheFile	Um local no Amazon S3 que contém os arquivos que o Hadoop vai copiar para o diretório de trabalho local (com o objetivo principal de melhorar a performance). Tipo: string Padrão: nenhum Restrições: [URI]#[nome do symlink a ser criado no diretório de trabalho]	Não
-cacheArchive	Arquivo JAR a ser extraído e colocado no diretório de trabalho Tipo: string Padrão: nenhum Restrições: [URI]#[nome do diretório do symlink a ser criado no diretório de trabalho]	Não
-combiner	Combina os resultados Tipo: string Padrão: nenhum Restrições: nome da classe Java	Não

O código de exemplo a seguir é um executável do mapeador escrito em Python. Esse script faz parte do aplicativo WordCount de amostra.

```
#!/usr/bin/python
import sys

def main(argv):
  line = sys.stdin.readline()
  try:
    while line:
      line = line.rstrip()
      words = line.split()
      for word in words:
        print "LongValueSum:" + word + "\t" + "1"
      line = sys.stdin.readline()
  except "end of file":
    return None
if __name__ == "__main__":
  main(sys.argv)
```

Enviar uma etapa de streaming

Esta seção aborda os conceitos básicos do envio de uma etapa de streaming para um cluster. Uma aplicação de streaming lê os dados de entrada de uma entrada padrão e, em seguida, executa um script ou executável (denominado mapeador) para cada entrada. Os resultados obtidos de cada entrada são salvos localmente, normalmente em uma partição do Hadoop Distributed File System (HDFS). Depois que todas as entradas são processadas pelo mapeador, um segundo script ou programa executável (chamado de reducer) processa os resultados do mapeador. Os resultados do reducer são enviados para a saída padrão. Você pode encadear uma série de etapas de streaming, em que a saída de uma etapa se torna a entrada de outra etapa.

O mapeador e o reducer podem ser referenciados como arquivos ou você pode fornecer uma classe Java. Você pode implementar o mapeador e o reducer em qualquer uma das linguagens compatíveis, incluindo Ruby, Perl, Python, PHP ou Bash.

Enviar uma etapa de streaming usando o console

Este exemplo descreve como usar o console do Amazon EMR para enviar uma etapa de streaming para um cluster em execução.

Enviar uma etapa de streaming

1. Abra o console do Amazon EMR em https://console.aws.amazon.com /emr.
2. Em Cluster List (Lista de clusters), selecione o nome do cluster.
3. Role até a seção Steps (Etapas) e expanda-a. Em seguida, escolha Add step (Adicionar etapa).
4. Na caixa de diálogo Add Step (Adicionar etapa):
   - Em Step type (Tipo de etapa), escolha Streaming program (Programa de streaming).
   - Em Name (Nome), aceite o nome padrão (programa Streaming) ou digite um novo nome.
   - Em Mapper, digite ou navegue até o local da classe do mapeador no Hadoop ou de um bucket do S3, onde se encontra o executável do mapeador como, por exemplo um programa Python. O valor do caminho deve estar no formato*BucketName*/*path*/*MapperExecutable*.
   - Em Reducer (Redutor), digite ou navegue até o local da classe do reducer no Hadoop ou de um bucket do S3 onde se encontra o executável do reducer, por exemplo um programa Python. O valor do caminho deve estar no formato*BucketName*/*path*/*MapperExecutable*. O Amazon EMR é compatível com a palavra-chave especial aggregate. Para obter mais informações, acesse a biblioteca Aggregate fornecida pelo Hadoop.
   - Em Input S3 location (Localização do S3 de entrada), digite ou navegue até o local dos dados de entrada.
   - Em Local de saída do S3, digite ou navegue até o nome do bucket de saída do Amazon S3.
   - Para Arguments (Argumentos), deixe o campo em branco.
   - Para Action on failure (Ação na falha), aceite a opção padrão Continue (Continuar).
5. Escolha Adicionar. A etapa é exibida no console com o status Pendente.
6. O status da etapa muda de Pending (Pendente) para Running (Em execução) e depois para Completed (Concluído) conforme ela é executada. Para atualizar o status, escolha o ícone Refresh (Atualizar) acima da coluna Actions (Ações).

AWS CLI

Esses exemplos demonstram como usar o AWS CLI para criar um cluster e enviar uma etapa de streaming.

Criar um cluster e enviar uma etapa de streaming usando a AWS CLI

- Para criar um cluster e enviar uma etapa de streaming usando a AWS CLI, digite o seguinte comando e *myKey* substitua pelo nome do seu EC2 key pair. Observe que o argumento para --

`files` deve ser o caminho do Amazon S3 para o local do script e os argumentos para `-mapper` e `-reducer` devem ser os nomes dos respectivos arquivos de script.

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.9.0 --
applications Name=Hue Name=Hive Name=Pig --use-default-roles \
--ec2-attributes KeyName=myKey --instance-type m5.xlarge --instance-count 3 \
--steps Type=STREAMING,Name="Streaming Program",ActionOnFailure=CONTINUE,Args=[--
files,pathtoscripts,-mapper,mapperscript,-reducer,reducerscript,aggregate,-
input,pathtoinputdata,-output,pathtooutputbucket]
```

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

Quando você especifica a contagem de instâncias sem usar o parâmetro `--instance-groups`, um único nó principal é executado, e as instâncias restantes são executadas como nós core. Todos os nós usam o tipo de instância especificado no comando.

> **Note**
>
> Se você não tiver criado o perfil de serviço padrão do Amazon EMR e o perfil de EC2 instância, digite aws `emr create-default-roles` para criá-los antes de digitar o subcomando. `create-cluster`

Para obter mais informações sobre o uso dos comandos do Amazon EMR no AWS CLI, consulte. https://docs.aws.amazon.com/cli/latest/reference/emr

Processar dados com um JAR personalizado

Um JAR personalizado executa um programa Java compilado que você pode carregar no Amazon S3. Você deve compilar o programa com base na versão do Hadoop que você deseja iniciar e enviar uma etapa `CUSTOM_JAR` ao cluster do Amazon EMR. Para obter mais informações sobre como compilar um arquivo JAR, consulte Criar binários usando o Amazon EMR.

Para obter mais informações sobre a criação de uma MapReduce aplicação do Hadoop, consulte o MapReduce Tutorial na documentação do Apache Hadoop.

Tópicos

- Enviar uma etapa de JAR personalizado

Enviar uma etapa de JAR personalizado

Um JAR personalizado executa um programa Java compilado que você pode carregar no Amazon S3. Você deve compilar o programa com base na versão do Hadoop que você deseja iniciar e enviar uma etapa `CUSTOM_JAR` ao cluster do Amazon EMR. Para obter mais informações sobre como compilar um arquivo JAR, consulte Criar binários usando o Amazon EMR.

Para obter mais informações sobre a criação de uma MapReduce aplicação do Hadoop, consulte o MapReduce Tutorial na documentação do Apache Hadoop.

Esta seção aborda os conceitos básicos do envio de uma etapa de JAR personalizado no Amazon EMR. O envio de uma etapa de JAR personalizado permite a gravação de um script para processar dados usando a linguagem de programação Java.

Enviar uma etapa de JAR personalizado usando o console

Este exemplo descreve como usar o console do Amazon EMR para enviar uma etapa de JAR personalizado a um cluster em execução.

Enviar uma etapa de JAR personalizado usando o console

1. Abra o console do Amazon EMR em https://console.aws.amazon.com /emr.
2. Em Cluster List (Lista de clusters), selecione o nome do cluster.
3. Role até a seção Steps (Etapas) e expanda-a. Em seguida, escolha Add step (Adicionar etapa).
4. Na caixa de diálogo Add Step (Adicionar etapa):
   - Para Step type (Tipo de etapa), escolha Custom JAR (JAR personalizado).
   - Para Name (Nome), aceite o nome padrão (JAR Personalizado) ou digite um novo nome.
   - Para JAR S3 location (Localização do S3 JAR), digite a localização do seu arquivo JAR ou navegue até ela. A localização do JAR pode ser um caminho para o S3 ou uma classe java totalmente qualificada no caminho de classe.

- Para Arguments (Argumentos), digite argumentos necessários como strings separadas por espaços ou deixe o campo em branco.
- Para Action on failure (Ação na falha), aceite a opção padrão Continue (Continuar).

5. Escolha Adicionar. A etapa é exibida no console com o status Pendente.
6. O status da etapa muda de Pending (Pendente) para Running (Em execução) e depois para Completed (Concluído) conforme ela é executada. Para atualizar o status, escolha o ícone Refresh (Atualizar) acima da coluna Actions (Ações).

Iniciar um cluster e enviar uma etapa de JAR personalizado usando a AWS CLI

Iniciar um cluster e enviar uma etapa de JAR personalizado usando a AWS CLI

Para iniciar um cluster e enviar uma etapa de JAR personalizado usando a AWS CLI, digite o `create-cluster` subcomando com o `--steps` parâmetro.

- Para iniciar um cluster e enviar uma etapa de JAR personalizado, digite o comando a seguir, *myKey* substitua pelo nome do seu EC2 key pair e *amzn-s3-demo-bucket* substitua pelo nome do seu bucket.

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.9.0 \
--applications Name=Hue Name=Hive Name=Pig --use-default-roles \
--ec2-attributes KeyName=myKey --instance-type m5.xlarge --instance-count 3 \
--steps Type=CUSTOM_JAR,Name="Custom JAR
 Step",ActionOnFailure=CONTINUE,Jar=pathtojarfile,Args=["pathtoinputdata","pathtooutputbuck
```

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

Quando você especifica a contagem de instâncias sem usar o parâmetro `--instance-groups`, um único nó primário é iniciado e as instâncias restantes são iniciadas como nós centrais. Todos os nós usam o tipo de instância que você especifica no comando.

> **Note**
>
> Se você não tiver criado o perfil de serviço padrão do Amazon EMR e o perfil de EC2 instância, digite `aws emr create-default-roles` para criá-los antes de digitar o subcomando. `create-cluster`

Para obter mais informações sobre o uso dos comandos do Amazon EMR no AWS CLI, consulte. https://docs.aws.amazon.com/cli/latest/reference/emr

Dependências de terceiros

Às vezes, pode ser necessário incluir no MapReduce classpath JARs para uso com seu programa. Existem duas opções para fazer isso:

- Inclua o `--libjars s3://URI_to_JAR` nas opções da etapa para o procedimento em Iniciar um cluster e enviar uma etapa de JAR personalizado usando a AWS CLI.
- Inicie o cluster com uma configuração `mapreduce.application.classpath` modificada em `mapred-site.xml`. Use a classificação para a configuração `mapred-site`. Para criar o cluster com a etapa usando a AWS CLI, a aparência seria a seguinte:

```
aws emr create-cluster --release-label emr-7.9.0 \
--applications Name=Hue Name=Hive Name=Pig --use-default-roles \
--instance-type m5.xlarge --instance-count 2  --ec2-attributes KeyName=myKey \
--steps Type=CUSTOM_JAR,Name="Custom JAR
 Step",ActionOnFailure=CONTINUE,Jar=pathtojarfile,Args=["pathtoinputdata","pathtooutputbucket
 \
--configurations https://s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myConfig.json
```

`myConfig.json`:

```
[
    {
      "Classification": "mapred-site",
      "Properties": {
        "mapreduce.application.classpath": "path1,path2"
      }
    }
```

```
]
```

A lista separada por vírgulas de caminhos deve ser acrescentada ao classpath da JVM de cada tarefa.

Ler objetos restaurados

Com o Amazon EMR versão 7.2.0 e superior, você pode ler objetos Glacier restaurados na localização do S3 da tabela com o protocolo `S3A`. Os mecanismos das versões anteriores não fazem distinção entre arquivos Glacier e Glacier Deep Archive, o que significa que você obteria um `AmazonS3Exception` se tentasse acessar um arquivo Glacier em andamento com `S3A`. Essa operação de leitura ignora os arquivos arquivados do Glacier se eles ainda estiverem sendo restaurados. Para habilitar esse comportamento, use a configuração `fs.s3a.glacier.read.restored.objects`. Essa configuração pode ter três valores:

- READ_ALL: esse valor indica que o Amazon EMR não deve contabilizar as classes de armazenamento recuperadas do Amazon S3. Esse é o comportamento padrão.
- SKIP_ALL_GLACIER: esse valor indica que o Amazon EMR deve ignorar qualquer objeto do S3 que esteja marcado com a classe de armazenamento Glacier e recuperar todos os outros objetos. Esse é o comportamento padrão do Amazon Athena em relação aos objetos Glacier.
- READ_RESTORED_GLACIER_OBJECTS: esse valor indica que o Amazon EMR deve verificar o status restaurado do objeto Glacier. Se o Amazon EMR puder restaurar o objeto, você poderá lê-lo como um objeto S3 normal. Caso contrário, o Amazon EMR ignorará o objeto Glacier do S3

Exemplos

Spark

Para ler objetos restaurados ao usar o Spark, use a seguinte configuração:

```
--conf spark.hadoop.fs.s3a.glacier.read.restored.objects=<value>
```

Se você usa spark-sql, utilize a seguinte configuração:

```
spark-sql --conf spark.hadoop.fs.s3a.glacier.read.restored.objects=<value>
```

Flink

Se você usa o Flink, pode definir a configuração no arquivo `flink-conf.yaml`

```
fs.s3a.glacier.read.restored.objects: <value>
```

Você também pode definir a classificação `flink-conf`:

```
[
 {
  "Classification": "flink-conf",
  "Properties": {
  "fs.s3a.glacier.read.restored.objects":"<value>"
  }
 }
]
```

Hive

Se você usa o Hive, defina a configuração no arquivo `hive-site.xml`.

```
<property>
 <name>fs.s3a.glacier.read.restored.objects</name>
 <value><value></value>
</property>
```

Você também pode usar a CLI do Hive para definir a propriedade `--hiveconf`:

```
hive --hiveconf fs.s3a.glacier.read.restored.objects=<value>
```

Considerações

Ao ler objetos Glacier restaurados do S3, observe as seguintes considerações:

- Você só pode ler objetos restaurados se usar o esquema `S3A` ou o `S3AFileSystem` para acessar os dados.
- Quando você lê um objeto Glacier restaurado, o Amazon EMR não restaura o objeto em si. Para fazer isso, você deve usar o AWS CLI ou o AWS SDK.

Ative o reconhecimento de acesso não uniforme à memória para contêineres do YARN

Com as versões 6.x e posteriores do Amazon EMR, você pode usar o acesso não uniforme à memória (NUMA) para multiprocessar seus dados em clusters. NUMA é um padrão de projeto de memória de computação em que o processador pode acessar sua própria memória local mais rapidamente do que a memória em outro processador ou compartilhada entre processadores. Os contêineres do YARN têm melhor performance com o NUMA porque podem se vincular a um nó NUMA específico que atende a todas as alocações de memória subsequentes. Isso reduz a quantidade de vezes que o cluster precisa acessar a memória remota.

É possível ativar o suporte do NUMA para o contêiner YARN quando a máquina do nó de processamento for um nó multi-NUMA. Para confirmar se o nó de processamento é um nó NUMA único ou multi-NUMA, execute o comando a seguir.

```
lscpu | grep -i numa
NUMA node(s): 2
```

Em geral, instâncias superiores a 12x têm dois nós NUMA. Isso não se aplica a instâncias metal.

Ativar o reconhecimento de NUMA para contêineres do YARN

1. Use a seguinte configuração `yarn-site` no cluster do Amazon EMR 6.x.

```
 [
 {
  "classification":"yarn-site",
   "properties":{
    "yarn.nodemanager.linux-container-executor.nonsecure-mode.local-user":"yarn",
    "yarn.nodemanager.linux-container-executor.group":"yarn",
    "yarn.nodemanager.container-
executor.class":"org.apache.hadoop.yarn.server.nodemanager.LinuxContainerExecutor",
    "yarn.nodemanager.numa-awareness.enabled":"true",
    "yarn.nodemanager.numa-awareness.numactl.cmd":"/usr/bin/numactl",
    "yarn.nodemanager.numa-awareness.read-topology":"true"
   },
  "configurations":[]
 }
 ]
```

2. Forneça a ação de bootstrap a seguir no cluster.

```
#!/bin/bash

sudo yum -y install numactl
echo 1 | sudo tee /proc/sys/kernel/numa_balancing

echo "banned.users=mapred,bin,hdfs" >> /etc/hadoop/conf/container-executor.cfg
rm -rf /var/log/hadoop-yarn/
sudo chown -R yarn:hadoop /var/log/hadoop-yarn/
sudo chmod 755 -R /var/log/hadoop-yarn/

sudo chmod 6050 /etc/hadoop/conf/container-executor.cfg

mkdir /mnt/yarn && sudo chmod 755 -R /mnt/yarn && sudo chown -R yarn:hadoop /mnt/
yarn
mkdir /mnt1/yarn && sudo chmod 755 -R /mnt1/yarn && sudo chown -R yarn:hadoop /
mnt1/yarn
mkdir /mnt2/yarn && sudo chmod 755 -R /mnt2/yarn && sudo chown -R yarn:hadoop /
mnt2/yarn
```

3. Cada contêiner deve ter reconhecimento de NUMA. É possível notificar a Máquina Virtual Java (JVM) em cada contêiner com um sinalizador NUMA. Por exemplo, para notificar a JVM para usar NUMA em um trabalho do MapReduce, adicione as propriedades a seguir em `mapred-site.xml`.

```
<property>
 <name>mapreduce.reduce.java.opts</name>
 <value>-XX:+UseNUMA</value>
</property>
<property>
 <name>mapreduce.map.java.opts</name>
 <value>-XX:+UseNUMA</value>
</property>
```

4. Para verificar se você ativou o NUMA, pesquise algum dos arquivos de log do NodeManager com o comando a seguir.

```
grep "NUMA resources allocation is enabled," *
```

Para verificar se NodeManager atribuiu recursos de NUMA nó a um contêiner, pesquise o NodeManager log com o comando a seguir, substituindo-o *<container_id>* pelo seu próprio ID de contêiner.

```
grep "NUMA node" | grep <container_id>
```

Embalagem de contêiner YARN

A partir da versão 7.9.0 do Amazon EMR, a política de empacotamento de contêineres agora está disponível para o programador de capacidade do YARN, que se baseia na política de posicionamento de vários nós do YARN. Embora o recurso esteja desativado por padrão, quando ativado, o YARN prioriza o preenchimento de um único nó com contêineres antes de expandir para outros nós do cluster, respeitando um limite de empacotamento predefinido definido pela configuração. `yarn.scheduler.capacity.multi-node-placement.container.bin-packing.percentage`

A política de embalagem de contêineres oferece vários benefícios em comparação com a estratégia padrão de alocação uniforme de contêineres:

- Reduz a fragmentação dos recursos do cluster.
- Ele potencialmente acelera as operações de redução de escala do cluster lançando contêineres em um número limitado de nós quando há recursos disponíveis nesses nós, deixando outros nós ociosos, que podem então ser reduzidos, resultando em uma melhor economia de custos para escalar dinamicamente um cluster.

Ative o recurso

Para habilitar o recurso de empacotamento de contêineres no Amazon EMR, você pode adicionar a seguinte classificação de site do YARN:

```
[
 {
  "Classification": "yarn-site",
  "Properties": {
  "yarn.scheduler.capacity.multi-node-placement.container.bin-packing.percentage":
 "integer value from 1-100"
  }
```

```
 }
]
```

Considerações

- O recurso está disponível exclusivamente para o programador de capacidade do YARN.
- A ativação do recurso ativa automaticamente a estratégia de agendamento de posicionamento de vários nós do YARN.
- Pode haver uma possível degradação do desempenho devido à utilização concentrada de recursos em um número limitado de nós.
- Com esse recurso, políticas personalizadas de auto-escalonamento demonstram melhores operações de redução, em comparação com a política de escalabilidade gerenciada.

Histórico de versões do Hadoop

A tabela a seguir informa a versão do Hadoop incluída em cada versão do Amazon EMR, junto com os componentes instalados com a aplicação. Para obter as versões dos componentes em cada versão, consulte a seção Versão de componentes da sua versão em Versões 7.x do Amazon EMR, Versões de lançamento 6.x do Amazon EMR ou Versões de lançamento 5.x do Amazon EMR.

Informações de versão do Hadoop

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-7.9.0	3.4.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-7.8.0	3.4.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.7.0	3.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-7.6.0	3.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.5.0	3.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-7.4.0	3.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.3.0	3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-7.2.0	3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.36.2	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-7.1.0	3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.0.0	3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-6.15.0	3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.14.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-6.13.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.12.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-6.11.1	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.11.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-6.10.1	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.10.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-6.9.1	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.9.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-6.8.1	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.8.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-6.7.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.36.1	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.36.0	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.6.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.35.0	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.5.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-6.4.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.3.1	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-6.3.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.2.1	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-6.2.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.1.1	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-6.1.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.0.1	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-6.0.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.34.0	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.33.1	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.33.0	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.32.1	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.32.0	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.31.1	2.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.31.0	2.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.30.2	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.30.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.30.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.29.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.28.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.28.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.27.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.27.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.26.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.25.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.24.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.24.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.23.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.23.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.22.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.21.2	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.21.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.21.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.20.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.20.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.19.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.19.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.18.1	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.18.0	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.17.2	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.17.1	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.17.0	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.16.1	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.16.0	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.15.1	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.15.0	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.14.2	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.14.1	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.14.0	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.13.1	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.13.0	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.12.3	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.12.2	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.12.1	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.12.0	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.11.4	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.11.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.11.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.11.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.11.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.10.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.10.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.9.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.9.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.8.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.8.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.8.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.8.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.7.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.7.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.6.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.6.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.5.4	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.5.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.5.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.5.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.5.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.4.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.4.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.3.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.3.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.3.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.2.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.2.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.2.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.2.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.1.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.1.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.0.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-5.0.2	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.0.1	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.0.0	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-4.9.6	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.9.5	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.9.4	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-4.9.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.9.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.9.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-4.8.5	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.8.4	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.8.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-4.8.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.8.1	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.8.0	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-4.7.4	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.7.3	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.7.2	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-4.7.1	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.7.0	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.6.1	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-4.6.0	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.5.0	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.4.0	2.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-4.3.0	2.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.2.0	2.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.1.0	2.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Rótulo de versão do Amazon EMR	Versão do Hadoop	Componentes instalados com o Hadoop
emr-4.0.0	2.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Notas de versão do Hadoop por versão

Tópicos

- Amazon EMR 7.9.0 - Notas de lançamento do Hadoop
- Amazon EMR 7.8.0 — Notas de lançamento do Hadoop
- Amazon EMR 7.7.0 — Notas de lançamento do Hadoop
- Amazon EMR 7.6.0 - Notas de lançamento do Hadoop
- Amazon EMR 7.5.0 - Notas de lançamento do Hadoop
- Amazon EMR 7.4.0 — Notas de lançamento do Hadoop
- Amazon EMR 7.3.0: notas de versão do Hadoop
- Amazon EMR 7.2.0: notas de versão do Hadoop
- Amazon EMR 6.6.0: notas de versão do Hadoop

Amazon EMR 7.9.0 - Notas de lançamento do Hadoop

Amazon EMR 7.9.0 - Alterações no Hadoop

Tipo	Descrição
Novo recurso	Mapeamento automático da configuração: a configuração do sistema de arquivos EMRFS

Tipo	Descrição
	é aplicada automaticamente à configuração correspondente do sistema de arquivos S3A, permitindo uma migração perfeita do EMRFS para o S3A.
Novo recurso	Embalagem de contêineres do YARN: política de agendamento para redução agressiva de escala
Backport	YARN-11752: Agendador global: melhore o tempo de alocação de contêineres
Correção de bugs	S3A: Fix IAMCredentials Provider devolvendo credenciais expiradas
Correção de bugs	Fixar valor de `yarn.log.server.url` configuração para clusters com criptografia em trânsito ativada

Amazon EMR 7.8.0 — Notas de lançamento do Hadoop

Amazon EMR 7.8.0 - Alterações no Hadoop

Tipo	Descrição
Upgrade	O Hadoop foi atualizado para a versão 3.4.1.

Amazon EMR 7.7.0 — Notas de lançamento do Hadoop

Amazon EMR 7.7.0 - Alterações no Hadoop

Tipo	Descrição
Novo recurso	Otimize a GlobStatus chamada S3A com a lista de prefixos do S3
Backport	YARN-7327: Habilite o agendamento assíncron o por padrão para o agendador de capacidade
Backport	YARN-10058: Lidar com exceções não detectadas para threads de agendamen to assíncrono para evitar travamentos do agendador
Backport	YARN-11732: Corrija o potencial NPE ao chamar #reservedContainer para Scheduler Node CapacityScheduler
Backport	YARN-11560: Corrija o bug do NPE quando vários nós estão habilitados com agendamento assíncrono
Backport	YARN-11191: Corrija um possível impasse em RefreshQueues GlobalScheduler
Backport	YARN-11041: Substitua todas as ocorrências de QueuePath pela nova classe QueuePath
Backport	YARN-11660: Corrige a regressão de desempenho para SingleConstraintAp pPlacementAllocator

Tipo	Descrição
Backport	HADOOP-19116: Atualização para o cliente zookeeper 3.8.4 devido ao CVE-2024-23944.
Backport	HADOOP-19115: Atualize para 9.37.2 devido ao nimbus-jose-jwt CVE-2023-52428.
Backport	HADOOP-19024: Use o bouncycastle jdk18 1.77
Backport	HADOOP-19123: Atualização para commons-configuration2 2.10.1 devido ao CVE
Backport	HADOOP-19114: Atualize para commons-compress 1.26.1 devido a CVEs
Backport	HADOOP-19237: Atualize para o dnsjava 3.6.1 devido a CVEs
Novo recurso	Adicionar auditoria de solicitações do S3 ao S3A
Backport	HADOOP-17609: Torne o suporte SM4 opcional para código nativo do OpenSSL
Backport	HADOOP-18583: o hadoop checknative falha ao carregar o openssl 3.x
Novo recurso	Adicionar suporte para mapeamentos de funções do S3A

Amazon EMR 7.7.0 - Recursos do Hadoop

- O agendamento assíncrono de contêineres se tornou a estratégia de agendamento padrão para o agendador de capacidade, projetado para otimizar a velocidade de alocação de contêineres.
- O sistema de arquivos S3A introduz uma otimização para chamadas de status global usando a listagem de prefixos do S3 para acelerar as operações da lista. Por padrão, esse recurso está desativado e pode ser ativado configurando `fs.s3a.prefix.listing.in.glob.status.enabled=true` no arquivo core-site.xml. Quando ativada, a otimização permite a filtragem do lado do servidor para chamadas globstatus`fs.globstatus("s3://bucket/a*")`, como melhorar o desempenho da lista listando somente os objetos que começam com. `"a"`
- Adicione a auditoria de solicitações do S3 ao S3A. Quando ativada, as informações do fileSystemOwner objeto são usadas para preencher a string userAgent com os campos do usuário e do grupo de usuários que fazem as solicitações do S3.
- O S3A adiciona suporte para mapeamentos de funções, o que ajuda a determinar qual função do IAM usar com base em usuários, grupos ou prefixos do S3.

Amazon EMR 7.6.0 - Notas de lançamento do Hadoop

Amazon EMR 7.6.0 - Alterações no Hadoop

Tipo	Descrição
Novo recurso	Estrutura da venda de credenciais S3A
Novo recurso	Suporte do S3A Magic Committer para Hive
Backport	HADOOP-19309: S3A: a CopyFromLocalFile operação falha quando o arquivo de origem não contém o esquema de arquivo
Backport	HADOOP-19050 S3A: suporte a concessões de acesso ao S3

Tipo	Descrição
Backport	HADOOP-18708 S3A: Support S3 Client Side Encryption (CSE)

Amazon EMR 7.6.0 - Recursos do Hadoop

- As configurações padrão para `yarn.scheduler.increment-allocation-mb` e `yarn.scheduler.minimum-allocation-mb` foram modificadas de 32 para 1, o que efetivamente desativa a normalização de recursos de contêiner no YARN.
- O sistema de arquivos S3A no Hadoop foi aprimorado para oferecer suporte à venda de credenciais em nível de solicitação. Com esse recurso, cada solicitação do Amazon S3 feita utilizará um. `AwsCredentialsProvider` Isso garante maior flexibilidade e segurança ao fornecer dinamicamente as credenciais para cada solicitação.
- O sistema de arquivos S3A no Hadoop oferece suporte à criptografia do lado do cliente S3.

Amazon EMR 7.5.0 - Notas de lançamento do Hadoop

Amazon EMR 7.5.0 - Alterações no Hadoop

Tipo	Descrição
Correção de bugs	Comentou fs.file.impl para um valor vazio.
Backport	HADOOP-19286: Support o acesso entre regiões do S3A quando o S3 está configurado region/endpoint
Melhoria	Configuração automática da região S3 para o conector S3A no EMR- EC2
Melhoria	Reduza o número de HeadObject chamadas no S3A

Tipo	Descrição

Com o lançamento do Amazon EMR 7.5, o conector S3A do Spark demonstra desempenho de leitura comparável ao EMRFS, conforme evidenciado por benchmarks usando um conjunto de dados de parquete TPC-DS de 3 TB.

Amazon EMR 7.5.0 - Recursos do Hadoop

- A configuração da região do S3 `fs.s3a.endpoint.region` é definida automaticamente para a região em que o cluster do EMR é iniciado com o conector S3A para implantação do EMR. EC2
- O acesso entre regiões entre buckets do Amazon S3 está habilitado por padrão para o conector S3A. Ele pode ser modificado por configuração`fs.s3a.cross.region.access.enabled=true or false`.

Amazon EMR 7.4.0 — Notas de lançamento do Hadoop

Amazon EMR 7.4.0 - Alterações no Hadoop

Tipo	Descrição
Upgrade	A versão do Hadoop foi atualizada para 3.4.0, consulte as notas de lançamento do OSS.
Correção de bugs	Corrija métricas negativas do Yarn pendentes e alocadas para FairScheduler
Correção de bugs	YARN-11702: Fixe o fio sobre a alocação de contêineres
Correção de bugs	Melhore o tratamento das condições de corrida ao reduzir a escala dos nós
Melhoria	HADOOP-18679: Adicionar API para exclusão de arquivos bulk/paged

Tipo	Descrição
Melhoria	HADOOP-19203: API WrappEdio para aumentar como não verificada BulkDelete IOEs IOExceptions
Melhoria	HADOOP-19205: S3A: mais lento do que com o SDK v1 initialization/close
Melhoria	HADOOP-19161: S3A: opção fs.s3a.performance.flags para obter uma lista de sinalizadores de desempenho
Melhoria	HADOOP-19072: S3A: expanda as otimizações nas lojas com fs.s3a.performance.flags para mkdir

Amazon EMR 7.4.0 - Recursos do Hadoop

Consulte a lista a seguir para ver os novos recursos do Hadoop no Amazon EMR 7.4.0.

- Os valores de configuração padrão foram ajustados para um desempenho ideal:
  - mapreduce.input.fileinputformat.list-status.num-threads=10 — Isso é acima de 1.
  - fs.s3a.block.size=64M — Isso é acima de 32M.
  - fs.s3a.multipart.size=128M — Isso é acima de 64M.
- Out-of-the-box otimizações de aprimoramento de desempenho para acelerar MapReduce trabalhos com o sistema de arquivos S3A.

Amazon EMR 7.3.0: notas de versão do Hadoop

Amazon EMR 7.3.0: alterações no Hadoop

Tipo	Descrição

Tipo	Descrição
Novo recurso	HADOOP-19197: Contexto de criptografia de suporte do S3A AWS KMS
Novo recurso	HADOOP-18980: remapeamento do provedor de credenciais do S3A: torne extensível
Melhoria	HADOOP-18808: LogExactlyOnce para adicionar um método debug ()

Amazon EMR 7.3.0: recursos do Hadoop

Consulte a lista a seguir para ver os novos recursos do Hadoop no Amazon EMR 7.3.0.

- Por padrão, clusters com criptografia em trânsito habilitada por meio da configuração de segurança executarão o TLS 1.3 para Hadoop YARN (Resource Manager, Node Manager, Timeline Server e WebApplicationProxy), HDFS (, NameNode DataNode, Journal Node and DFSZKFailover Controller) e Map Reduce (portas Server e Shuffle). JobHistory MapReduce

Amazon EMR 7.2.0: notas de versão do Hadoop

Amazon EMR 7.2.0: alterações no Hadoop

Tipo	Descrição
Novo recurso	HADOOP-18850: S3A: habilite a criptografia de camada dupla do lado do servidor com chaves AWS KMS
Novo recurso	Suporte a objetos de leitura Glacier restaurados do S3
Melhoria	

Tipo	Descrição
	Ajuste de escala dinâmico das conexões máximas do S3A com base nos núcleos ao executar trabalhos do Spark
Melhoria	HADOOP-18797: suporte a gravações simultâneas com o S3A Magic Committer
Melhoria	HADOOP-19047: InMemory Support Rastreamento de Compromissos Magic do S3A
Melhoria	Adicione o método “Exportar”, “Publicar” e "AbortPublish" no MagicV2 Committer
Melhoria	Implemente a ReadFullyIntoBuffers FileSystem API no S3A
Correção de bugs	HADOOP-18793: O S3A não limpa o diretório de staging-uploads StagingCommitter
Upgrade	HADOOP-19008: S3A: atualize a versão do aws-sdk para 2.21.41
Upgrade	HADOOP-18613: Atualize para a versão 3.8.3 ZooKeeper

Amazon EMR 7.2.0: recursos do Hadoop

Confira a lista a seguir dos novos recursos do Hadoop no Amazon EMR 7.2.0.

- Suporte ao Glacier do S3 para ler objetos restaurados: com o Amazon EMR versão 7.2.0 e superior, você pode ler objetos Glacier restaurados da localização da tabela no S3 com o protocolo

`S3A`. Essa operação de leitura ignora os arquivos arquivados do Glacier se eles ainda estiverem sendo restaurados.

- Suporte a gravações simultâneas com o S3A Magic Committer: o Amazon EMR 7.2.0 introduz suporte para gravações simultâneas com o S3A Magic Committer. Esse recurso foi adicionado à comunidade OSS do Hadoop como parte do HADOOP-18797.

Amazon EMR 6.6.0: notas de versão do Hadoop

Amazon EMR 6.6.0: alterações no Hadoop

Tipo	Descrição
Bug	Registros duplicados corrigidos ao ler arquivos BZip2 de texto.
Backport	HADOOP-18136: verificar o tratamento de arquivos .tar ausentes por `FileUtils.unTar()`
Backport	HADOOP-17627: backport para branch-3.2 HADOOP-17371, HADOOP-17621, HADOOP-17625 para atualizar o Jetty para 9.4.39
Backport	HADOOP-17655: atualizar o Jetty para 9.4.40
Backport	HADOOP-17796: atualizar a versão do Jetty para 9.4.43
Backport	HADOOP-17661: mvn versions:set falha ao analisar pom.xml
Backport	HADOOP-17236: aumentar o snakeyaml para 1.26 para mitigar o CVE-2017-18640
Backport	HADOOP-16717: Remova a dependência de IsLog4JLogger no Log4 GenericsUtil jLoggerAdapter

Tipo	Descrição
Backport	HADOOP-17633: aumente o json-smart para 2.4.2 e 9.8 devido a nimbus-jose-jwt CVEs
Backport	HADOOP-17844: atualizar o JSON smart para 2.4.7
Backport	HADOOP-17972: backport do HADOOP-17 683 (atualização commons-io para 2.8.0) para branch-3.2
Backport	HADOOP-16555: atualizar o commons-c ompress para 1.19
Backport	HADOOP-17370: atualizar o commons-c ompress para 1.21
Backport	HADOOP-17096: Corrigir o deslocamento do buffer de entrada do compressor ZStandard
Backport	HADOOP-17112: não são permitidos espaços em branco nos caminhos ao salvar arquivos no s3a via confirmador
Backport	HADOOP-13500: sincronizar a iteração do objeto de propriedades de configuração
Backport	HDFS-14099: Descritor de quadro desconhec ido ao descompactar vários quadros no Descompressor ZStandard
Backport	HDFS-16410: Análise insegura de Xml em OfflineEditsXmlLoader
Backport	HDFS-14498: LeaseManager pode repetir para sempre o arquivo cuja criação falhou

Tipo	Descrição
Backport	HDFS-15290: entrada NPE durante a inicializ ação HttpServer NameNode
Backport	HDFS-15293: flexibilizar a condição para aceitar uma fsimage ao receber um ponto de verificação
Backport	HDFS-12979: StandbyNode deve ser carregado após o ponto de verificação FsImage ObserverNode
Backport	YARN-10538: adicionar nós de reativação à lista de nós atualizados retornados ao AM
Backport	YARN-10472: Backport YARN-10314 (lançado apenas com jars de cliente sombreados) para o YarnClient NoClassDefFoundError branch-3.2 WebSocketException
Backport	YARN-9968: o localizador público está saindo em `NodeManager` devido a `NullPoint erException`
Backport	YARN-10651: `CapacityScheduler` travou com NPE em `AbstractYarnSchedu ler.updateNodeResource()`
Backport	YARN-9339: métrica de aplicações pendentes incorreta após mover a aplicação para uma nova fila
Backport	YARN-10438: manipular `containerId` nulo em `ClientRMService#getContaine rReport()`

Tipo	Descrição
Backport	YARN-7266: o ATS 1.5 não conseguirá iniciar se os arquivos `RollingLevelDb` estiverem corrompidos ou ausentes
Backport	YARN-9063: o ATS 1.5 não conseguirá iniciar se os arquivos `RollingLevelDb` estiverem corrompidos ou ausentes
Backport	YARN-9848: reverter o YARN-4946 (o RM não deve considerar uma aplicação como `COMPLETED` quando a agregação de logs não está em um estado terminal).

Apache HBase

HBaseé um banco de dados distribuído, não relacional e de código aberto desenvolvido como parte do projeto Hadoop da Apache Software Foundation. HBase é executado sobre o Hadoop Distributed File System (HDFS) para fornecer recursos de banco de dados não relacional para o ecossistema Hadoop. HBase está incluído na versão 4.6.0 e posterior do Amazon EMR.

HBase funciona perfeitamente com o Hadoop, compartilhando seu sistema de arquivos e servindo como entrada e saída diretas para a MapReduce estrutura e o mecanismo de execução. HBase também se integra ao Apache Hive, permitindo consultas semelhantes a SQL em tabelas, uniões com HBase tabelas baseadas em Hive e suporte para conectividade de banco de dados Java (JDBC). Para obter mais informações sobre HBase, consulte o Apache HBase e a HBase documentação no site do Apache. Para ver um exemplo de como usar HBase com o Hive, consulte a postagem do blog sobre AWS Big Data Combine NoSQL e análises paralelas massivas usando o Apache e o HBase Apache Hive no Amazon EMR.

Com o HBase Amazon EMR, você também pode fazer backup de seus HBase dados diretamente no Amazon Simple Storage Service (Amazon S3) e restaurar a partir de um backup criado anteriormente ao iniciar um cluster. HBase O Amazon EMR oferece opções de integração adicionais com o Amazon S3 para recuperação de desastres e persistência de dados.

- HBase no Amazon S3 - Com o Amazon EMR versão 5.2.0 e posterior, você pode usar no Amazon HBase S3 para armazenar o diretório HBase raiz e os metadados de um cluster diretamente no Amazon S3. Posteriormente, você pode iniciar um novo cluster, apontando-o para o local do diretório de raiz no Amazon S3. Somente um cluster por vez pode usar a HBase localização no Amazon S3, com exceção de um cluster de réplica de leitura. Para obter mais informações, consulte HBase no Amazon S3 (modo de armazenamento Amazon S3).
- HBase réplicas de leitura - O Amazon EMR versão 5.7.0 e posterior com o Amazon HBase S3 oferece suporte a clusters de réplica de leitura. Um cluster de réplica de leitura fornece acesso somente leitura aos arquivos de armazenamento de um cluster primário e a metadados para operações somente leitura. Para obter mais informações, consulte Usar um cluster de réplica de leitura.
- HBase Snapshots - Como alternativa ao HBase Amazon S3, com o EMR versão 4.0 e posterior, você pode criar instantâneos de seus dados HBase diretamente no Amazon S3 e depois recuperar dados usando os snapshots. Para obter mais informações, consulte Usando HBase instantâneos.

> **Important**
>
> Para a escalabilidade de HBase clusters do Amazon EMR, não recomendamos o uso de escalabilidade gerenciada ou escalabilidade com políticas personalizadas com clusters. HBase

A tabela a seguir lista a versão HBase incluída na versão mais recente da série 7.x do Amazon EMR, junto com os componentes com os quais o Amazon EMR é instalado. HBase

Para a versão dos componentes instalados HBase nesta versão, consulte Versões de componentes da versão 7.9.0.

HBase informações da versão do emr-7.9.0

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-7.9.0	HBase 2.6.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

A tabela a seguir lista a versão HBase incluída na versão mais recente da série 6.x do Amazon EMR, junto com os componentes com os quais o Amazon EMR é instalado. HBase

Para a versão dos componentes instalados HBase nesta versão, consulte Versões de componentes da versão 6.15.0.

HBase informações da versão do emr-6.15.0

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-6.15.0	HBase 2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

> **Note**
>
> O Apache HBase HBCK2 é uma ferramenta operacional separada para reparar HBase regiões e tabelas do sistema. No Amazon EMR 6.1.0 e versões posteriores, o hbase-hbck2.jar é fornecido em `/usr/lib/hbase-operator-tools/` no nó primário. Para obter mais informações sobre como criar e usar a ferramenta, consulte HBaseHBCK2.

A tabela a seguir lista a versão HBase incluída na versão mais recente da série 5.x do Amazon EMR, junto com os componentes com os quais o Amazon EMR é instalado. HBase

Para a versão dos componentes instalados HBase nesta versão, consulte Versões de componentes da versão 5.36.2.

HBase informações da versão do emr-5.36.2

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.36.2	HBase 1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Tópicos

- Criação de um cluster com HBase
- HBase no Amazon S3 (modo de armazenamento Amazon S3)
- Logs de gravação antecipada (WAL) do Amazon EMR
- Usando a HBase concha
- Acesse HBase tabelas com o Hive
- Usando HBase instantâneos
- Configure HBase
- Visualize a interface HBase do usuário
- Exibir arquivos HBase de log
- Monitor HBase com Ganglia

- Monitorando o EMR com HBase a Amazon CloudWatch
- Migrando de versões anteriores HBase
- HBase histórico de lançamentos

Criação de um cluster com HBase

Os procedimentos desta seção abordam os princípios básicos da inicialização de um cluster usando o AWS Management Console e o. AWS CLI Para obter informações detalhadas sobre como planejar, configurar e executar clusters do Amazon EMR, consulte Plan and configure clusters no Guia de gerenciamento do Amazon EMR.

Criação de um cluster HBase usando o console

Para obter etapas rápidas para iniciar clusters com o console, consulte Getting started with Amazon EMR no Guia de gerenciamento do Amazon EMR.

Para iniciar um cluster com HBase instalado usando o console

1. Abra o console do Amazon EMR em https://console.aws.amazon.com /emr.
2. Escolha Create cluster (Criar cluster) e Go to advanced options (Ir para opções avançadas).
3. Para Software Configuration (Configuração do software), escolha um valor de 4.6.0 ou posterior para Amazon Release Version (Versão da Amazon)(recomendamos a versão mais recente). Escolha HBaseoutros aplicativos conforme desejado.
4. Com o Amazon EMR versão 5.2.0 e posterior, em Configurações de HBase armazenamento, selecione HDFS ou S3. Para obter mais informações, consulte HBase no Amazon S3 (modo de armazenamento Amazon S3).
5. Selecione outras opções conforme necessário e escolha Create cluster (Criar cluster).

Criando um cluster HBase usando o AWS CLI

Use o comando a seguir para criar um cluster com HBase instalado:

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.9.0 \
--applications Name=HBase --use-default-roles --ec2-attributes KeyName=myKey \
--instance-type m5.xlarge --instance-count 3
```

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

Se você usa HBase no Amazon S3, especifique a `--configurations` opção com uma referência a um objeto de configuração JSON. O objeto de configuração deve conter uma `hbase-site` classificação que especifique o local no Amazon S3 HBase onde os dados são armazenados usando `hbase.rootdir` a propriedade. Ele também deve conter uma classificação `hbase`, que especifica `s3` usando a propriedade `hbase.emr.storageMode`. O exemplo a seguir demonstra um trecho de código JSON com essas definições de configuração.

```
[
    {
        "Classification": "hbase-site",
        "Properties": {
            "hbase.rootdir": "s3://amzn-s3-demo-bucket/MyHBaseStore"
        }
    },
    {
        "Classification": "hbase",
        "Properties": {
            "hbase.emr.storageMode": "s3"
        }
    }
]
```

Para obter mais informações HBase sobre o Amazon S3, consulte. HBase no Amazon S3 (modo de armazenamento Amazon S3) Para obter mais informações sobre classificações, consulte Configurar aplicações.

HBase no Amazon S3 (modo de armazenamento Amazon S3)

Ao executar HBase no Amazon EMR versão 5.2.0 ou posterior, você pode habilitar no Amazon HBase S3, que oferece as seguintes vantagens:

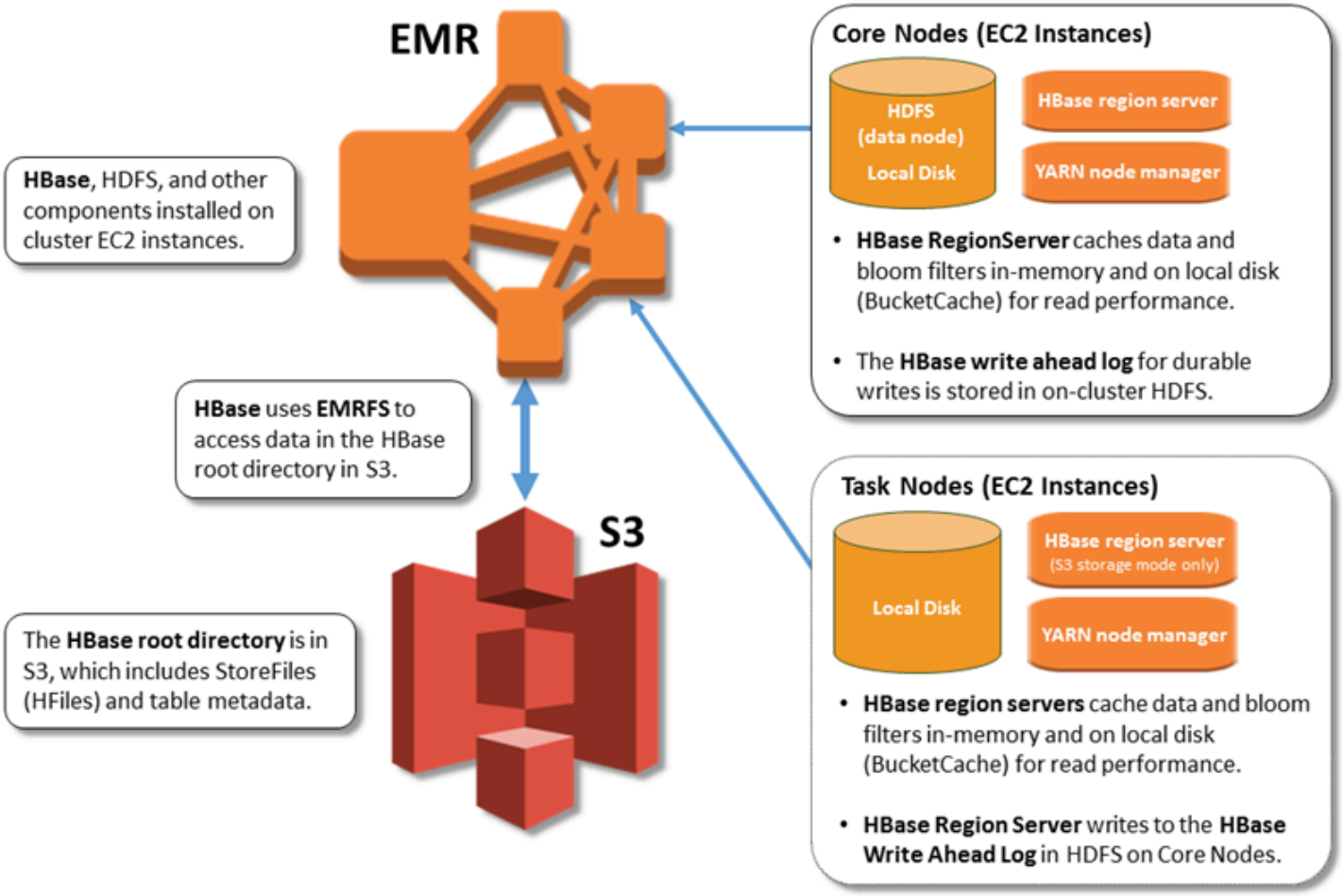
- O diretório HBase raiz é armazenado no Amazon S3, incluindo arquivos de HBase armazenamento e metadados de tabelas. Esses dados são persistentes fora do cluster,

disponíveis nas zonas de EC2 disponibilidade da Amazon, e você não precisa se recuperar usando snapshots ou outros métodos.

- Com arquivos de armazenamento no Amazon S3, você pode dimensionar seu cluster do Amazon EMR para suas necessidades de computação em vez de requisitos de dados, com replicação de 3x no HDFS.
- Usando o Amazon EMR versão 5.7.0 ou posterior, você pode configurar um cluster de réplica de leitura, que permite manter cópias somente leitura dos dados no Amazon S3. Você pode acessar os dados do cluster de réplica de leitura para realizar operações de leitura simultaneamente e no evento de o cluster primário se tornar indisponível.
- No Amazon EMR versão 6.2.0 a 7.3.0, o HFile rastreamento persistente usa uma tabela HBase do sistema chamada `hbase:storefile` para rastrear diretamente os HFile caminhos usados para operações de leitura. Esse atributo é habilitado por padrão e não exige que a migração manual seja executada. Nas versões superiores à 7.3.0, HFile os caminhos são rastreados usando um rastreador de arquivos, armazenando HFile os caminhos diretamente em um meta-arquivo, dentro do diretório de armazenamento.

> **Note**
>
> Usuários que estão usando uma versão do Amazon EMR anterior à 7.4.0 e estão migrando para o EMR-7.4.0 e posterior, consulte Migração de HBase versões anteriores e siga a documentação de atualização disponível para garantir uma transição tranquila.

A ilustração a seguir mostra os HBase componentes relevantes HBase no Amazon S3.



Habilitando HBase no Amazon S3

Você pode habilitar HBase no Amazon S3 usando o console do Amazon EMR, o AWS CLI ou a API do Amazon EMR. A configuração é uma opção durante a criação do cluster. Ao usar o console, você escolhe a configuração usando Advanced options (Opções avançadas). Quando você usa a AWS CLI, use a opção `--configurations` para fornecer um objeto de configuração JSON. As propriedades do objeto de configuração especificam o modo de armazenamento e o local do diretório raiz no Amazon S3. O local do Amazon S3 que você especificar deve estar na mesma região que seu cluster do Amazon EMR. Somente um cluster ativo por vez pode usar o mesmo diretório HBase raiz no Amazon S3. Para ver as etapas do console e um exemplo detalhado de criação de cluster usando o AWS CLI, consulte. Criação de um cluster com HBase Um objeto de configuração de exemplo é mostrado no seguinte trecho de código JSON.

```
{
  "Classification": "hbase-site",
```

```
  "Properties": {
    "hbase.rootdir": "s3://amzn-s3-demo-bucket/my-hbase-rootdir"}
},
{
  "Classification": "hbase",
  "Properties": {
  "hbase.emr.storageMode":"s3"
  }
}
```

> **Note**
>
> Se você usar um bucket do Amazon S3 como forma HBase, deverá adicionar uma barra no final do URI do Amazon S3. `rootdir` Por exemplo, é necessário usar `"hbase.rootdir: s3://amzn-s3-demo-bucket/"`, em vez de `"hbase.rootdir: s3://amzn-s3-demo-bucket"`, para evitar problemas.

Usar um cluster de réplica de leitura

Depois de configurar um cluster primário usando o HBase Amazon S3, você pode criar e configurar um cluster de réplica de leitura que fornece acesso somente de leitura aos mesmos dados do cluster primário. Isso é útil quando você precisa de acesso simultâneo para consultar dados ou de acesso ininterrupto caso o cluster primário se torne indisponível. O atributo de réplica de leitura está disponível no Amazon EMR 5.7.0 e versões posteriores.

O cluster primário e o cluster de réplica de leitura são configuradas da mesma maneira, com uma diferença importante. Ambos apontam para o mesmo local `hbase.rootdir`. No entanto, a classificação `hbase` para o cluster de réplica de leitura inclui a propriedade `"hbase.emr.readreplica.enabled":"true"`.

O cluster de réplica de leitura foi projetado para operações somente de leitura, e nenhuma ação manual de compactação ou gravação deve ser executada nele. Para versões do Amazon EMR anteriores à 7.4.0, é recomendável desativar a compactação no cluster de réplica de leitura ao ativar o recurso de réplica de leitura. Essa precaução é necessária porque, com o recurso de HFile rastreamento persistente ativado no cluster primário, é possível que o cluster de réplica de leitura compacte as tabelas do sistema, o que pode causar uma falha FileNotFoundException no cluster

primário. A desativação da compactação no cluster de réplica de leitura evita inconsistências de dados entre os clusters primário e de réplica de leitura.

Por exemplo, dada a classificação JSON para o cluster primário, conforme mostrado anteriormente neste tópico, a configuração de um cluster de réplica de leitura para versões do EMR anteriores à 7.4.0 é a seguinte:

```
{
  "Classification": "hbase-site",
  "Properties": {
    "hbase.rootdir": "s3://amzn-s3-demo-bucket/my-hbase-rootdir",
    "hbase.regionserver.compaction.enabled": "false"
  }
},
{
  "Classification": "hbase",
  "Properties": {
  "hbase.emr.storageMode":"s3",
  "hbase.emr.readreplica.enabled":"true"
  }
}
```

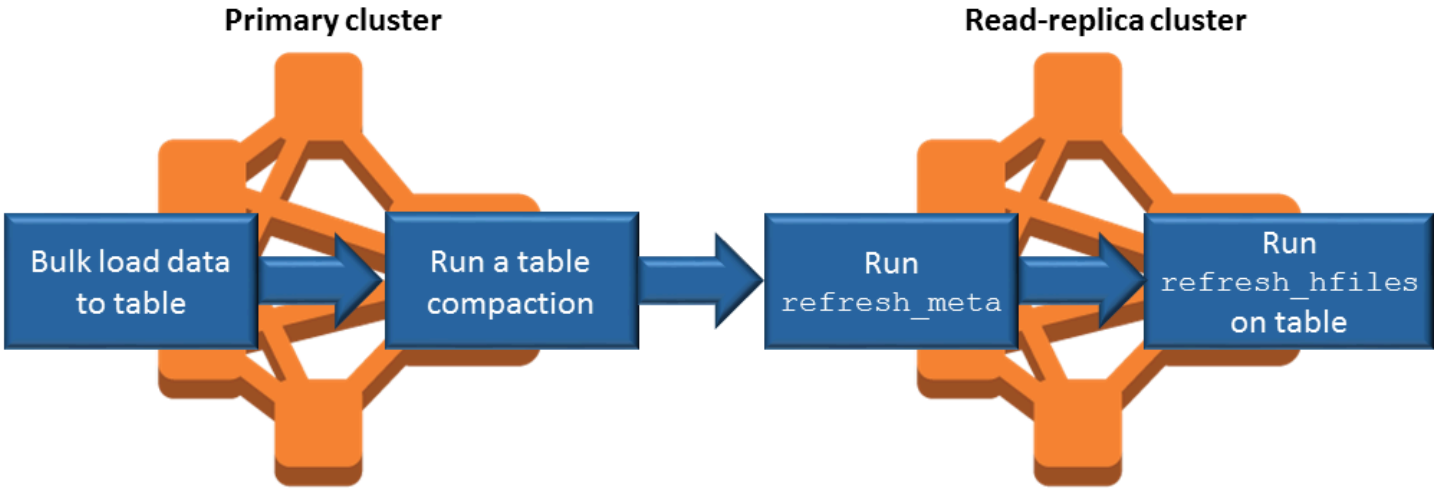
Para versões do Amazon EMR posteriores à 7.3.0, agora usamos o Acompanhamento de arquivos da loja recurso, portanto, não há necessidade de desativar as compactações.

Sincronizar a réplica de leitura ao adicionar dados

Como a réplica de leitura usa HBase StoreFiles e os metadados que o cluster primário grava no Amazon S3, a réplica de leitura é tão atual quanto o armazenamento de dados do Amazon S3. A orientação a seguir pode ajudar a minimizar o tempo de retardo entre o cluster primário e a réplica de leitura quando você grava dados.

- Carregue os dados em massa no cluster primário sempre que possível. Para obter mais informações, consulte Carregamento em massa na HBase documentação do Apache.
- Uma liberação que grava arquivos de armazenamento no Amazon S3 deve ocorrer logo que possível após a adição dos dados. Faça a liberação manualmente ou ajuste configurações de liberação para minimizar o tempo de retardo.

- Se compactações puderem ser executadas automaticamente, execute uma compactação manual para evitar inconsistências quando compactações forem acionadas.
- No cluster de réplica de leitura, quando algum metadado for alterado, por exemplo, quando ocorrerem divisões ou compactações de HBase regiões, ou quando tabelas forem adicionadas ou removidas, execute o comando. `refresh_meta`
- No cluster de réplica de leitura, execute o comando `refresh_hfiles` quando registros forem adicionados ou alterados em uma tabela.



HFile Monitoramento persistente

O HFile rastreamento persistente usa uma tabela HBase do sistema chamada `hbase:storefile` para rastrear diretamente os HFile caminhos usados para operações de leitura. Novos HFile caminhos são adicionados à tabela à medida que dados adicionais são adicionados HBase. Isso remove as operações de renomeação como um mecanismo de confirmação nas HBase operações críticas do caminho de gravação e melhora o tempo de recuperação ao abrir uma HBase região lendo a tabela do `hbase:storefile` sistema em vez da listagem de diretórios do sistema de arquivos. Esse recurso é ativado por padrão nas versões 6.2.0 a 7.3.0 do Amazon EMR e não exige nenhuma etapa de migração manual.

> **Note**
>
> O HFile rastreamento persistente usando a tabela do sistema HBase storefile não oferece suporte ao recurso de replicação da HBase região. Para obter mais informações sobre a

replicação HBase da região, consulte Leituras de alta disponibilidade consistentes com a linha do tempo.

Desativando o rastreamento persistente HFile

O HFile rastreamento persistente é ativado por padrão a partir da versão 6.2.0 do Amazon EMR. Para desativar o HFile rastreamento persistente, especifique a seguinte substituição de configuração ao iniciar um cluster:

```
{
  "Classification": "hbase-site",
  "Properties": {
    "hbase.storefile.tracking.persist.enabled":"false",

 "hbase.hstore.engine.class":"org.apache.hadoop.hbase.regionserver.DefaultStoreEngine"
  }
}
```

Note

Ao reconfigurar o cluster do Amazon EMR, todos os grupos de instâncias deverão ser atualizados.

Sincronizar a tabela Storefile manualmente

A tabela do storefile é mantida atualizada à medida que novas HFile instâncias são criadas. No entanto, se a tabela de arquivos de armazenamento ficar fora de sincronia com os arquivos de dados por qualquer motivo, os comandos a seguir poderão ser usados para sincronizar manualmente os dados:

Sincronizar a tabela storefile em uma região on-line:

```
hbase org.apache.hadoop.hbase.client.example.RefreshHFilesClient <table>
```

Sincronizar a tabela do storefile em uma região off-line:

- Remova o znode da tabela storefile.

```
echo "ls /hbase/storefile/loaded" | sudo -u hbase hbase zkcli
```

```
[<tableName>, hbase:namespace]
# The TableName exists in the list
echo "delete /hbase/storefile/loaded/<tableName>" | sudo -u hbase hbase zkcli
# Delete the Table ZNode
echo "ls /hbase/storefile/loaded" | sudo -u hbase hbase zkcli
[hbase:namespace]
```

- Atribua a região (execute em “hbase shell”).

```
hbase cli> assign '<region name>'
```

- Se a tarefa falhar.

```
hbase cli> disable '<table name>'
hbase cli> enable '<table name>'
```

Escalar a tabela storefile

Por padrão, a tabela storefile é dividida em quatro regiões. Se a tabela storefile continuar sob carga de gravação pesada, a tabela ainda poderá ser dividida manualmente.

Para dividir uma região ativa específica, use o comando a seguir (executado em “hbase shell”).

```
hbase cli> split '<region name>'
```

Para dividir a tabela, use o comando a seguir (execute em “hbase shell”).

```
hbase cli> split 'hbase:storefile'
```

Acompanhamento de arquivos da loja

Por padrão, usamos a FileBasedStoreFileTrackerimplementação. Essa implementação cria novos arquivos diretamente no diretório da loja, evitando a necessidade de operações de renomeação. Ele mantém uma lista de instâncias de hfile confirmadas na memória, apoiada por meta-arquivos em cada diretório de armazenamento. Sempre que um novo arquivo hfile é confirmado, a lista de arquivos rastreados no armazenamento fornecido é atualizada e um novo meta-arquivo é gravado com o conteúdo da lista e descartando o meta-arquivo anterior, que contém uma lista desatualizada. Mais informações sobre o rastreamento de arquivos da loja podem ser encontradas em Rastreamento de arquivos da loja no guia de HBase referência do Apache.

A implementação do FileBasedStoreFile rastreador é ativada por padrão, começando com a versão 7.4.0 do Amazon EMR:

```
{
  "Classification": "hbase-site",
  "Properties": {
    hbase.store.file-tracker.impl:
 "org.apache.hadoop.hbase.regionserver.storefiletracker.FileBasedStoreFileTracker"
}
```

Para desativar a FileBasedStoreFileTracker implementação, especifique a seguinte substituição de configuração ao iniciar um cluster:

```
{
  "Classification": "hbase-site",
  "Properties": {
    hbase.store.file-tracker.impl:
 "org.apache.hadoop.hbase.regionserver.storefiletracker.DefaultStoreFileTracker"
}
```

> **Note**
>
> Ao reconfigurar o cluster do Amazon EMR, todos os grupos de instâncias deverão ser atualizados.

Considerações operacionais

HBase os servidores regionais são usados BlockCache para armazenar leituras de dados na memória e BucketCache armazenar leituras de dados no disco local. Além disso, os servidores regionais usam MemStore para armazenar gravações de dados na memória e usam registros de gravação antecipada para armazenar gravações de dados no HDFS antes que os dados sejam gravados no Amazon HBase StoreFiles S3. O desempenho de leitura do seu cluster estão relacionado à frequência com a qual um registro pode ser recuperado dos caches na memória ou no disco. Uma perda de cache faz com que o registro seja lido StoreFile no Amazon S3, que tem latência significativamente maior e maior desvio padrão do que a leitura do HDFS. Além disso, as taxas máximas de solicitações para o Amazon S3 são menores do que as que podem ser obtidas no cache local e, portanto, o armazenamento de dados no cache pode ser importante para workload

com uso intensivo de leitura. Para obter mais informações sobre a performance do Amazon S3, consulte Performance optimization, no Guia do usuário do Amazon Simple Storage Service.

Para melhorar o desempenho, recomendamos que você armazene em cache o máximo possível do seu conjunto de dados no armazenamento da EC2 instância. Como BucketCache usa o armazenamento de EC2 instâncias do servidor da região, você pode escolher um tipo de EC2 instância com um armazenamento de instâncias suficiente e adicionar armazenamento do Amazon EBS para acomodar o tamanho de cache necessário. Você também pode aumentar o BucketCache tamanho dos armazenamentos de instâncias anexados e dos volumes do EBS usando a `hbase.bucketcache.size` propriedade. A configuração padrão é 8.192 MB.

Para gravações, a frequência de MemStore descargas e o número de StoreFiles presentes durante compactações menores e maiores podem contribuir significativamente para um aumento nos tempos de resposta do servidor regional. Para um desempenho ideal, considere aumentar o tamanho do multiplicador de MemStore fluxo e HRegion bloco, o que aumenta o tempo decorrido entre as principais compactações, mas também aumenta o atraso na consistência se você usar uma réplica de leitura. Em alguns casos, você pode obter melhor performance usando tamanhos de blocos de arquivos maiores (porém inferiores a 5 GB) para acionar a funcionalidade do carregamento multiparte do Amazon S3 no EMRFS. O tamanho padrão do bloco do Amazon EMR é 128 MB. Para obter mais informações, consulte Configuração do HDFS. Raramente há clientes que excedem 1 GB de tamanho de bloco ao fazer a comparação do desempenho com liberações e compactações. Além disso, HBase as compactações e os servidores regionais funcionam de maneira ideal quando menos StoreFiles precisam ser compactados.

Tabelas podem demorar um tempo significativo para serem descartadas no Amazon S3, pois diretórios grandes precisam ser renomeados. Considere desabilitar tabelas em vez de as descartar.

Existe um processo de HBase limpeza que limpa arquivos WAL antigos e armazena arquivos. Com o Amazon EMR versão 5.17.0 e posterior, o agente de limpeza está habilitado globalmente, e as seguintes propriedades de configuração podem ser usadas para controlar o comportamento do agente de limpeza.

Propriedade de configuração	Valor padrão	Descrição
`hbase.regionserver.hfilecleaner.large.thread.count`	1	O número de segmentos alocados para limpar expirou muito. HFiles

Propriedade de configuração	Valor padrão	Descrição
`hbase.regionserver.hfilecleaner.small.thread.count`	1	O número de segmentos alocados para limpar expirou pouco. HFiles
`hbase.cleaner.scan.dir.concurrent.size`	Defina como um quarto de todos os núcleos disponíveis.	O número de threads para verificar os WALs diretórios antigos.
`hbase.oldwals.cleaner.thread.size`	2	O número de threads a serem limpos no WALs WALs diretório antigo.

Com o Amazon EMR 5.17.0 e versões anteriores, a operação do agente de limpeza pode afetar a performance de consultas ao executar workloads pesadas. Por isso, recomendamos que você ative o agente de limpeza apenas fora de horários de pico. O limpador tem os seguintes comandos de HBase shell:

- `cleaner_chore_enabled` consulta se o agente de limpeza está habilitado.
- `cleaner_chore_run` executa manualmente o agente de limpeza para remover arquivos.
- `cleaner_chore_switch` habilita ou desabilita o agente de limpeza e retorna ao seu estado anterior. Por exemplo, `cleaner_chore_switch true` habilita o agente de limpeza.

Propriedades para ajuste HBase de desempenho no Amazon S3

Os parâmetros a seguir podem ser ajustados para ajustar o desempenho da sua carga de trabalho quando você usa HBase no Amazon S3.

Propriedade de configuração	Valor padrão	Descrição
`hbase.bucketcache.size`	8,192	A quantidade de espaço em disco, em MB, reservada nos armazenamentos de EC2 instâncias da Amazon

Propriedade de configuração	Valor padrão	Descrição
		no servidor regional e nos volumes do EBS para BucketCache armazenamento. A configuração se aplica a todas as instâncias do servidor de regiões. BucketCache Tamanhos maiores geralmente correspondem a um melhor desempenho
`hbase.hregion.memstore.flush.size`	134217728	O limite de dados, em bytes, no qual uma liberação de memstore para o Amazon S3 é acionada.
`hbase.hregion.memstore.block.multiplier`	4	Um multiplicador que determina o limite MemStore superior no qual as atualizações são bloqueadas. Se os MemStore excedentes forem `hbase.hregion.memstore.flush.size` multiplicados por esse valor, as atualizações serão bloqueadas. MemStore descargas e compactação podem ocorrer para desbloquear as atualizações.
`hbase.hstore.blockingStoreFiles`	10	O número máximo StoreFiles que pode existir em uma loja antes que as atualizações sejam bloqueadas.

Propriedade de configuração	Valor padrão	Descrição
`hbase.hregion.max.filesize`	10737418240	O tamanho máximo de uma região antes que ela seja dividida.

Desativar e restaurar um cluster sem perda de dados

Para desligar um cluster do Amazon EMR sem perder dados que não foram gravados no Amazon S3, você deve liberar seu cache no Amazon S3 MemStore para gravar novos arquivos de armazenamento. Primeiro, é necessário desabilitar todas as tabelas. A seguinte configuração de etapas pode ser usada quando você adiciona uma etapa ao cluster. Para obter mais informações, consulte Work with steps using the AWS CLI and console no Guia de gerenciamento do Amazon EMR.

```
Name="Disable all tables",Jar="command-runner.jar",Args=["/bin/bash","/usr/lib/hbase/
bin/disable_all_tables.sh"]
```

Como alternativa, você pode executar o seguinte comando bash diretamente.

```
bash /usr/lib/hbase/bin/disable_all_tables.sh
```

Depois de desativar todas as tabelas, limpe a `hbase:meta` tabela usando o HBase shell e o comando a seguir.

```
flush 'hbase:meta'
```

Em seguida, você pode executar um script de shell fornecido no cluster do Amazon EMR para limpar o cache. MemStore Você pode adicioná-lo como uma etapa ou executá-lo diretamente, usando a AWS CLI no cluster. O script desativa todas as HBase tabelas, o que faz com que o servidor MemStore de cada região seja transferido para o Amazon S3. Se o script for concluído com êxito, os dados persistirão no Amazon S3, e o cluster poderá ser terminado.

Para reiniciar um cluster com os mesmos HBase dados, especifique a mesma localização do Amazon S3 do cluster anterior na propriedade de configuração AWS Management Console ou usando a propriedade de `hbase.rootdir` configuração.

Logs de gravação antecipada (WAL) do Amazon EMR

Com o Amazon EMR 6.15 e versões posteriores, você pode gravar seus registros de gravação HBase antecipada (WAL) do Apache no Amazon EMR WAL. Com versões mais baixas do Amazon EMR, quando você cria um cluster com a opção HBase no Amazon S3, o WAL é o único componente do HBase Apache que é armazenado no disco local para clusters, e você pode armazenar outros componentes, como o diretório raiz, armazenar arquivos HFiles (), metadados de tabelas e dados no Amazon S3.

Você pode usar o Amazon EMR WAL para recuperar dados que não foram descarregados no Amazon S3. Para fazer backup completo de seus HBase clusters, opte por usar o serviço WAL do Amazon EMR. Nos bastidores, `RegionServer` grava seus registros HBase de gravação antecipada (WAL) no WAL do Amazon EMR.

Caso seu cluster ou a AZ não estejam íntegros ou estejam indisponíveis, você pode criar um cluster, direcioná-lo para o mesmo diretório raiz do S3 e Workspace do Amazon EMR WAL e recuperar automaticamente os dados no WAL em alguns minutos. Para obter mais informações, consulte Restauração do Amazon EMR WAL.

Começando com as versões 7.3.0 e superiores do Amazon EMR, o Amazon EMR cria vários EMR para cada servidor e agrupa várias regiões HBase em um Amazon WALs EMR WAL. Isso aprimora o Apache HBase WAL para melhorar a utilização dos registros e otimizar os custos. Para configurar o número de instâncias WAL do Amazon EMR por HBase `RegionServer`, use o parâmetro. `hbase.wal.regiongrouping.numgroups` Por padrão, esse parâmetro é definido como 2. Existem duas tabelas de sistema que não estão incluídas em nenhum grupo WAL: meta e masterstore. Essas tabelas sempre usam seu próprio indivíduo WALs.

Se você executar uma versão inferior ao Amazon EMR 7.3.0, recomendamos que você desabilite manualmente as tabelas no HBase cluster antigo para garantir que todos os dados no Amazon EMR WAL sejam liberados para o Amazon S3. Em seguida, exclua o Amazon EMR WAL antigo, encerre o cluster antigo e configure um novo cluster que execute a versão mais recente. Se você tiver problemas e não conseguir desabilitar as tabelas no cluster antigo, poderá encerrar diretamente o cluster antigo e definir `emr.wal.multiplex.migrate` como `true` no novo cluster. Se definido como verdadeiro, HBase tentará reproduzir os dados de instâncias WAL antigas do Amazon EMR HBase durante a inicialização da região e excluir os antigos após a repetição. WALs Esse processo de reprodução gera custos adicionais para leituras. Após a migração, recomendamos configurar o cluster e definir `emr.wal.multiplex.migrate` como `false`. Como alternativa, você pode remover o parâmetro para acelerar a inicialização HBase da região.

> **Note**
>
> O Amazon EMR WAL exclui os dados depois de liberá-los. HBase Se HBase não liberar os dados, o Amazon EMR WAL retém os dados por no máximo 30 dias. Depois de 30 dias, o Amazon EMR WAL exclui automaticamente os dados. O Amazon EMR mantém instâncias do WAL por até 30 dias a partir do encerramento de um cluster do EMR. No entanto, se você iniciar um novo cluster habilitado para WAL no mesmo diretório raiz do S3 dentro desses 30 dias, o Amazon EMR não excluirá nenhuma das instâncias do WAL do cluster anterior. Para obter mais informações, consulte Restauração do Amazon EMR WAL.

As seções a seguir descrevem como configurar e usar o Amazon EMR WAL com seu cluster EMR habilitado HBase.

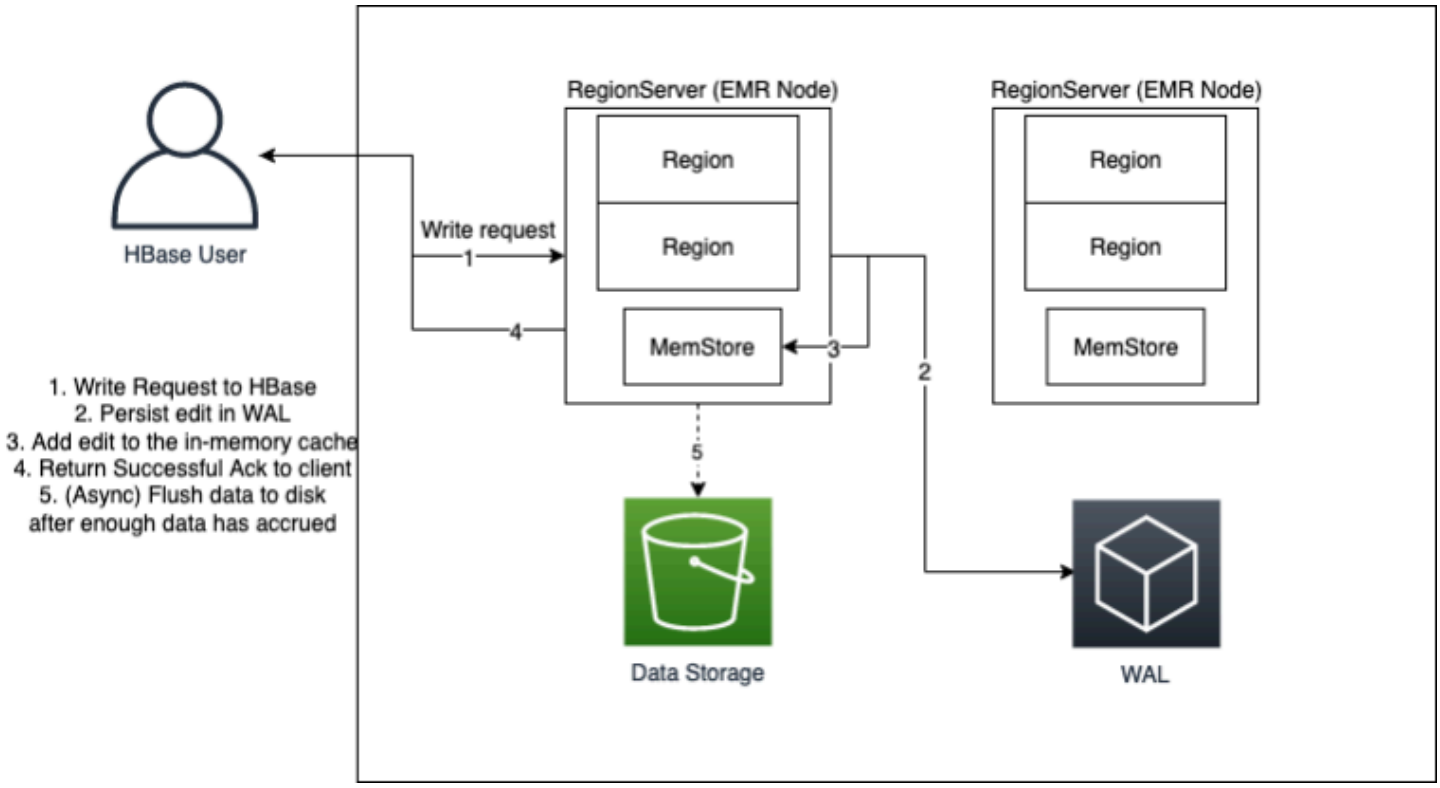
Tópicos

- Workspaces do Amazon EMR WAL
- Permissões necessárias para o Amazon EMR WAL
- Como habilitar o Amazon EMR WAL
- Restauração do Amazon EMR WAL
- Uso das configurações de segurança com o Amazon EMR WAL
- Acesso ao Amazon EMR WAL por meio do AWS PrivateLink
- Noções básicas de preços e métricas do Amazon EMR WAL
- Marcação de Workspaces do WAL
- Replicação entre clusters do EMR WAL
- Considerações e regiões do Amazon EMR WAL
- Referência da CLI do Amazon EMR WAL (EMRWAL)

Workspaces do Amazon EMR WAL

O Amazon EMR WAL adiciona o conceito de Workspaces do WAL. Um espaço de trabalho WAL é um contêiner lógico de. WALs Cada log de gravação antecipada no Amazon EMR WAL é encapsulado por um Workspace do WAL. Um cluster EMR grava WALs em exatamente um espaço de trabalho WAL que você configura na inicialização do cluster ou no, `defaultWALworkspace` se você não especificar um espaço de trabalho. Os espaços de trabalho do WAL não estão relacionados a nenhuma HBase terminologia existente, como namespaces.

Você pode usar os Workspaces do WAL para definir o escopo das permissões do IAM relativas ao Amazon EMR WAL para incluir somente os Workspaces que o cluster precisa acessar. Você também pode marcar o Workspace do WAL para controle de acesso baseado em tags. Para obter mais informações sobre tags, consulte Marcação de Workspaces do WAL.



Permissões necessárias para o Amazon EMR WAL

Para que o cluster se conecte ao Amazon EMR WAL, o perfil de instância do cluster requer determinadas permissões do IAM:

- O Amazon EMR WAL usa o perfil AWSServiceRoleForEMRWAL vinculado ao serviço para recuperar o status de um cluster. O Amazon EMR cria automaticamente essa função vinculada ao serviço quando você cria um espaço de trabalho do WAL, ou cria a função vinculada ao serviço quando HBase você configura um espaço de trabalho para o Amazon EMR WAL e a função vinculada ao serviço ainda não existe.

  Antes de habilitar o Amazon EMR WAL para um cluster, você deve configurar as permissões para permitir a criação automática do perfil AWSServiceRoleForEMRWAL vinculado ao serviço. Para obter mais informações e um exemplo de instrução que adiciona esse recurso, consulte Using service-linked roles for write-ahead logging.

- Como o Amazon EMR WAL usa o HBase Write Ahead Log (WAL), seus clusters devem usar o WAL. HBase A seguir estão as permissões mínimas do IAM que você precisa executar HBase. Adicione-os à política de permissões do seu perfil de instância:

```
emrwal:DeleteWal
emrwal:CreateWal
emrwal:CreateWorkspace
emrwal:AppendEdit
emrwal:ReplayEdits
emrwal:GetCurrentWalTime
emrwal:CompleteWalFlush
emrwal:ListWALs
emrwal:DescribeWAL
emrwal:TrimWAL
emrwal:ArchiveWAL
emrwal:ArchiveWALCheckPoint
```

> **Note**
>
> Se você definir o escopo das permissões do Amazon EMR WAL apenas para o conjunto mínimo, alguns comandos da CLI do EMRWAL não terão as permissões necessárias para serem executados.

Como habilitar o Amazon EMR WAL

Use as etapas a seguir para habilitar a gravação no Amazon EMR WAL ao criar um cluster com a AWS Command Line Interface.

> **Note**
>
> Você não pode habilitar o Amazon EMR WAL para um cluster que já está em execução e iniciar dois clusters com o mesmo diretório raiz do S3. Para obter mais informações, consulte Considerações e regiões do Amazon EMR WAL.

1. Antes de criar um cluster habilitado para Amazon EMR WAL, você deve adicionar as permissões necessárias ao perfil de instância que planeja usar com o cluster. Para obter mais informações, consulte Permissões necessárias para o Amazon EMR WAL.

2. Crie um cluster com a AWS CLI. Use a opção `--configurations` para fornecer um objeto de configuração JSON que especifique a propriedade `hbase.emr.wal.enabled`, conforme mostrado no exemplo abaixo.

   - Especifique o modo de armazenamento e o local do diretório raiz no Amazon S3. A localização do Amazon S3 que você especificar deve estar na mesma região do seu cluster EMR, mas somente um cluster ativo pode usar o mesmo diretório HBase raiz no S3 por vez.
   - Crie um cluster com a configuração de grupos de instâncias. Você não pode usar o Amazon EMR WAL com a configuração de frotas de instâncias. Para obter mais informações sobre a criação de clusters com grupos de instâncias, consulte Configure uniform instance groups no Guia de gerenciamento do Amazon EMR.
   - Para ver as etapas do console para criar um cluster e um `create-cluster` exemplo detalhado que usa o AWS CLI, consulte Como criar um cluster com HBase.

3. Para habilitar o WAL no novo cluster, defina a propriedade `hbase.emr.wal.enabled` como `true`. O comando a seguir contém um trecho JSON com um exemplo de objeto de configuração.

```
aws emr create-cluster --name "hbasewal" --release-label emr-6.x.y \
--applications Name=HBase --use-default-roles --ec2-attributes KeyName=myKey \
--instance-type m6i.xlarge --instance-count 1 --configurations hbase.json
$cat hbase.json
[
    {
        "Classification": "hbase-site",
        "Properties": {
            "hbase.rootdir": "s3://amzn-s3-demo-bucket/MyHBaseStore"
        }
    },
    {
        "Classification": "hbase",
        "Properties": {
            "hbase.emr.storageMode": "s3",
            "hbase.emr.wal.enabled": "true"
        }
    }
]
```

Quando HBase estiver on-line no cluster recém-criado, HBase gravará automaticamente os dados do WAL no Amazon EMR WAL e usará o Amazon EMR WAL para fins de recuperação.

Example 1: Criação de um cluster do EMR que use o Amazon EMR WAL

```
[
    {
        "Classification": "hbase-site",
        "Properties": {
            "hbase.rootdir": "s3://amzn-s3-demo-bucket/MyHBaseStore"
        }
    },
    {
        "Classification": "hbase",
        "Properties": {
            "hbase.emr.storageMode": "s3",
            "hbase.emr.wal.enabled": "true"
        }
    }
]
```

Example 2: Criação de um cluster do EMR com um Workspacede do WAL personalizado

```
[
    {
        "Classification": "hbase-site",
        "Properties": {
            "hbase.rootdir": "s3://amzn-s3-demo-bucket/MyHBaseStore",
            "emr.wal.workspace": "customWorkspaceName"
        }
    },
    {
        "Classification": "hbase",
        "Properties": {
            "hbase.emr.storageMode": "s3",
            "hbase.emr.wal.enabled": "true"
        }
    }
]
```

Restauração do Amazon EMR WAL

Como o Amazon EMR WAL do cluster original é retido por 30 dias, você pode restaurar e reutilizar o WAL para um cluster recém-criado dentro desse período. Quando você executa um novo cluster

no mesmo diretório raiz do S3, o Amazon EMR mantém as instâncias do WAL do cluster antigo. Se você encerrar esse novo cluster, o relógio de 30 dias será reiniciado a partir do encerramento.

Use o procedimento a seguir para restaurar um WAL existente com um novo cluster. Esse processo pressupõe que você criou o cluster original com o Amazon EMR WAL habilitado.

1. Dentro de 30 dias após a criação de um cluster habilitado para WAL, crie um novo cluster Região da AWS igual ao cluster original. O novo cluster pode estar na mesma AZ ou em uma AZ diferente dentro da mesma região em que o cluster original foi criado.

   Configure as propriedades do objeto para especificar o modo de armazenamento e o local do diretório raiz no Amazon S3. A localização do Amazon S3 que você especificar deve estar na mesma região do seu cluster EMR, mas somente um cluster ativo pode usar o mesmo diretório HBase raiz no S3 por vez.

   Para ver as etapas do console para criar um cluster e um `create-cluster` exemplo detalhado que usa o AWS CLI, consulte Como criar um cluster com HBase.
2. Para usar o Amazon EMR WAL existente no novo cluster, defina a propriedade `hbase.emr.wal.enabled` como `true`. O trecho JSON a seguir mostra um exemplo de objeto de configuração.

```
[
    {
        "Classification": "hbase-site",
        "Properties": {
            "hbase.rootdir": "s3://amzn-s3-demo-bucket/MyHBaseStore"
        }
    },
    {
        "Classification": "hbase",
        "Properties": {
            "hbase.emr.storageMode": "s3",
            "hbase.emr.wal.enabled": "true"
        }
    }
]
```

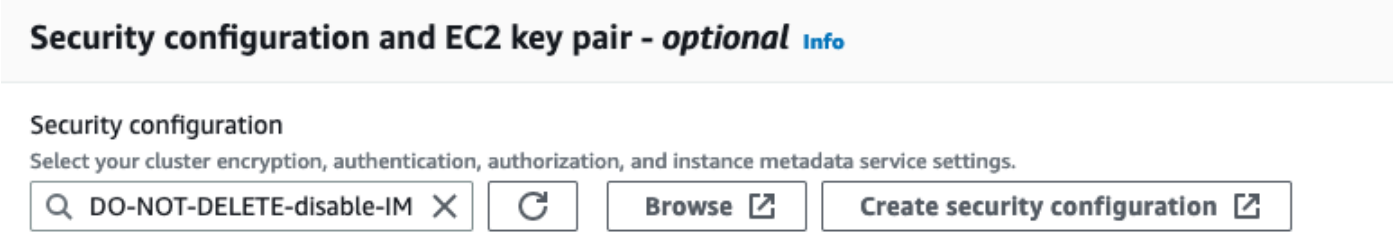
Uso das configurações de segurança com o Amazon EMR WAL

O Amazon EMR criptografa automaticamente os dados em trânsito entre o cluster e o serviço Amazon EMR WAL, bem como os dados em repouso no Amazon EMR WAL. Para obter mais informações, consulte Encryption at rest for Amazon EMR WAL. Você também pode usar uma configuração de segurança para trazer suas próprias chaves do serviço AWS Key Management Service (KMS) e criptografar os dados que você armazena no Amazon EMR WAL.

Use um dos métodos a seguir para selecionar uma configuração de segurança ao criar um cluster:

Console

Em AWS Management Console, especifique a configuração em Configuração de segurança e par de EC2 chaves.



CLI

A partir do AWS CLI, defina o `--security-configuration` parâmetro ao usar o comando create-cluster.

Para obter mais informações, consulte Encryption at rest for Amazon EMR WAL e Use security configurations to set up cluster security no Guia de gerenciamento do Amazon EMR.

Para obter mais informações relacionadas à segurança sobre o WAL, consulte Using service-linked roles for write-ahead logging.

Acesso ao Amazon EMR WAL por meio do AWS PrivateLink

Se você quiser manter sua conexão dentro da AWS rede, o Amazon EMR WAL oferece suporte. AWS PrivateLink Para configurar AWS PrivateLink, use o AWS Management Console ou AWS Command Line Interface (AWS CLI) para criar uma interface VPC endpoint que se conecta ao Amazon EMR WAL. Para obter mais informações, consulte Acessar um AWS serviço usando uma interface VPC endpoint no Guia.AWS PrivateLink

As etapas básicas são:

1. Use o console da Amazon VPC para criar um endpoint da VPC. Selecione Endpoints e Criar endpoint.
2. Mantenha a categoria Serviço como Serviços da AWS .
3. Na barra de pesquisa do painel Serviços, digite **`emrwal`** e selecione o serviço rotulado `com.amazonaws.`*`region`*`.emrwal.prod`.
4. Selecione sua VPC e salve o endpoint. Certifique-se de anexar ao endpoint da VPC os mesmos grupos de segurança que você anexa ao cluster do EMR.
5. Se quiser, agora você pode habilitar nomes de host DNS privados para o novo endpoint. Defina Habilitar nomes de host DNS e Habilitar suporte de DNS como `true` para a VPC. Em seguida, selecione o ID de endpoint, escolha Editar configurações de VPC no menu Ações e habilite nomes DNS privados.

   - Os nomes de host DNS privados do endpoint seguirão o formato `prod.emrwal.`*`region`*`.amazonaws.com`.
   - Se você não habilitar nomes de host DNS privados, a Amazon VPC fornecerá um nome de endpoint do DNS no formato *`endpointID`*`.prod.emrwal.`*`region`*`.vpce.amazonaws.com`.
6. Para usar seu AWS PrivateLink endpoint, modifique a `emr.wal.client.endpoint` configuração ao criar seu cluster habilitado para Amazon EMR WAL, conforme mostrado no exemplo a seguir:

```
[
    {
        "Classification": "hbase-site",
        "Properties": {
            "hbase.rootdir": "s3://amzn-s3-demo-bucket/MyHBaseStore",
            "emr.wal.workspace": "customWorkspaceName",
            "emr.wal.client.endpoint": "https://prod.emrwal.region.amazonaws.com"
        }
    },
    {
        "Classification": "hbase",
        "Properties": {
            "hbase.emr.storageMode": "s3",
            "hbase.emr.wal.enabled": "true"
        }
    }
]
```

Você também pode usar a política de VPCE para permitir ou restringir o acesso ao Amazon EMR WAL. APIs Para obter mais informações, consulte Controlar o acesso aos endpoints da VPC usando políticas de endpoint no Guia AWS PrivateLink .

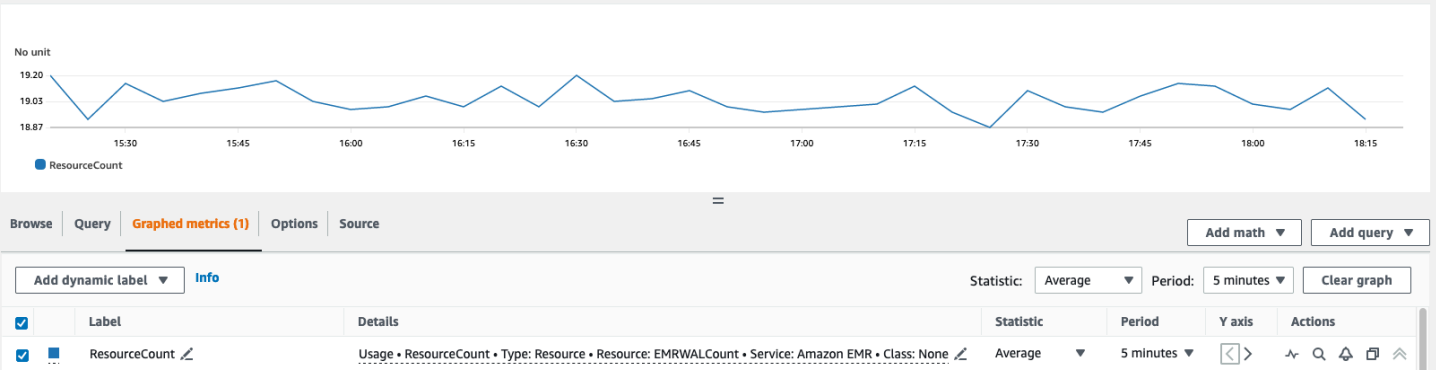
Noções básicas de preços e métricas do Amazon EMR WAL

Unidade de cobrança de recursos centrais	Detalhes
EMR-WAL-Read-GiB	API calls to read data from your table are billed as ReadRequestGiB. This includes `Get` and `Scan` operations. Reads are charged based on the sizes of the read items. Amazon EMR bills at a minimum of 1 byte. For example, if you read a 1234.12 bytes item, you're charged for 1235 bytes. Reads are aggregated every hour for billing and shown as GiBs.
EMR-WAL-Write-GiB	API calls to write data from your table are billed as Write-GiB. This includes `Put` operations. Writes are charged based on the sizes of the written items. Amazon EMR bills at a minimum of 1 byte. For example, if you write a 1234.12 bytes item, you're charged for 1235 bytes. Writes are aggregated every hour for billing and shown as GiBs.

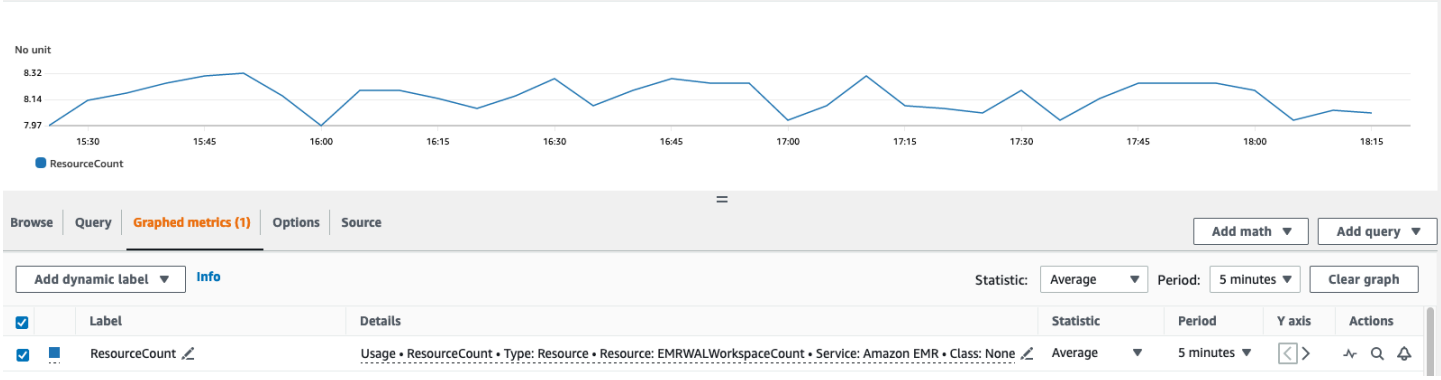
EMR-WAL-WALHours	O número WALs que você armazena no serviço é cobrado como`EMR-WAL-WALHours` . O Amazon EMR cria um WAL por região. HBase Por exemplo, se você criar 20 HBase tabelas, incluindo tabelas do sistema, e cada tabela tiver duas HBase regiões, você usará 28.800 horas de WAL, calculadas como:
	`20 tables`
	`x 2 Regions per table`
	`x 1 WAL per Region`
	`x 30 days`
	`x 24 hours`
	`----------`

Unidade de cobrança de recursos centrais	Detalhes
	`28,800 EMR-WAL-WALHours`

Exemplo**`EMRWALCount`**:



Exemplo**`EMRWALWorkspaceCount`**:



Marcação de Workspaces do WAL

Você pode adicionar tags a um Workspace ao criar um, além de adicionar, remover ou listar tags de um Workspace ativo para um cluster em execução. Você não pode marcar os recursos individuais no Workspace e nem atualizar as tags existentes; em vez disso, remova as tags indesejadas do Workspace e substitua-as.

Você pode marcar Workspaces na EMRWAL CLI. Para obter uma lista de comandos da EMRWAL CLI para marcar Workspaces, consulte Referência da CLI do Amazon EMR WAL (EMRWAL).

O seguinte exemplo de política do IAM ilustra um cenário que permite operações CRUDL no Workspace somente com a chave `resource_tag_allow_test_key` e o valor `resource_tag_allow_test_value` de marcação adequados:

```
{
  "Version": "2012-10-17",
  "Statement": [
      {
          "Action": [
              "EMRWAL:*"
          ],
          "Effect": "Allow",
          "Resource": [
              "*",
              "*"
          ],
          "Condition": {
              "StringEquals": {
                  "aws:ResourceTag/resource_tag_allow_test_key": [
                      "resource_tag_allow_test_value"
                  ]
              }
          }
      }
  ]
}
```

Para verificar se a tag agora é necessária nas operações do Workspace, use Referência da CLI do Amazon EMR WAL (EMRWAL) para chamar o comando `listTagsForResource` em `tagAllowResourceTag` para o Workspace com a tag de recurso desejada. Se você configurou a condição corretamente, o comando será bem-sucedido.

```
emrwal listTagsForResource -r us-east-1 -arn arn:aws:emrwal:us-east-1:arn:workspace/
tagAllowResourceTag
Tag(Key=resource_tag_allow_test_key, Value=resource_tag_allow_test_value)
```

Replicação entre clusters do EMR WAL

A partir do EMR 7.5, o EMR WAL oferece suporte à replicação HBase entre clusters de registros de gravação antecipada. Este tópico mostra como ativar o recurso e verificar se ele está funcionando.

Para obter mais detalhes sobre a replicação em cluster, consulte Replicação em [cluster na documentação do Apache.](#) HBase

> **Note**
>
> Há um custo extra de leitura associado aos registros de gravação antecipada, porque o processo de replicação lê dados do EMR WAL local. Para obter mais detalhes sobre o custo, consulte [Sobre os lançamentos do Amazon EMR.](#)

Configuração de replicação entre clusters

Para ter a mesma experiência do usuário, habilitar o recurso de replicação no EMR WAL é o mesmo que com HBase registros de gravação antecipada nativos. O procedimento a seguir mostra um exemplo simples. Para obter mais informações, consulte a HBase documentação em [Cluster Replication.](#)

1. Inicie um cluster primário, que é a fonte de replicação, com o EMR WAL ativado. Para habilitar registros de gravação antecipada, consulte. [Como habilitar o Amazon EMR WAL](#) Além disso, inicie um cluster de pares. Para esse cluster de mesmo nível, você pode optar por habilitar o EMR WAL ou não.
2. Nos dois clusters, crie uma tabela:

```
HBASE_CMD="sudo -u hbase hbase"
echo "create 'test_replication_table',{NAME => 'CF'}" | $HBASE_CMD shell
```

3. Adicione a configuração de pares no cluster primário e habilite a replicação de tabelas. Durante a adição de pares, ele precisa do nome de host do nó principal do cluster de pares, que é PEER_DNS.

```
HBASE_CMD="sudo -u hbase hbase"
PEER_DNS="ip-10-1-1-0.ec2.com"
PEER_NAME="aws"
TABLE_NAME="test_replication_table"

## Create peering with the destination cluster
echo "add_peer '$PEER_NAME', CLUSTER_KEY => '$PEER_DNS:2181:/hbase'" | $HBASE_CMD
 shell

## List peers in the primary cluster to confirm peer setup
```

```
echo "list_peers" | $HBASE_CMD shell

## Enable table replication
echo "enable_table_replication '$TABLE_NAME'" | $HBASE_CMD shell
```

Confirmando a replicação entre clusters

Depois de executar as etapas de configuração, a replicação é habilitada entre o cluster primário e o cluster de mesmo nível. Segue um teste que confirma que a replicação está funcionando.

1. Adicione dados no cluster primário e verifique os dados replicados no cluster de mesmo nível.

```
## Write on primary cluster with HBase CLI

put 'test_replication_table', 'aaa', 'CF:a', 'aaa_a1'
put 'test_replication_table', 'bbb', 'CF:b', 'bbb_b1'
put 'test_replication_table', 'ccc', 'CF:c', 'ccc_c1'
```

2. Confirme a replicação bem-sucedida no cluster de mesmo nível. Nesse caso, você deve ver dados replicados gravados do cluster primário para o cluster de mesmo nível.

```
### Scan on peer cluster with HBase CLI

scan 'test_replication_table'
```

Considerações e regiões do Amazon EMR WAL

Considerações sobre o Amazon EMR WAL

A seguinte lista descreve considerações e limitações importantes do Amazon EMR WAL:

- O Amazon EMR WAL está disponível para uso com as versões 6.15.0 e posteriores do Amazon EMR.
- O Amazon EMR WAL é um serviço pago opcional. Você paga pelo que usa: leituras, gravações e armazenamento de dados. Para obter mais informações, consulte Noções básicas de preços e métricas do Amazon EMR WAL e a página Preços do Amazon EMR.
- O Amazon EMR WAL usa o HBase Write Ahead Log (WAL). Para usar o Amazon EMR WAL, seus clusters devem usar o WAL. HBase

- Para habilitar o Amazon EMR WAL ao criar um cluster, você deve ter as permissões de perfil necessárias. Para obter mais informações, consulte Using service-linked roles for write-ahead logging.
- Você deve habilitar o Amazon EMR WAL ao criar o cluster com a API,, ou AWS Management Console AWS CLI, e você deve usar a configuração de grupos de instâncias. Você não pode habilitar o Amazon EMR WAL em um cluster em execução se não tiver criado o cluster com o Amazon EMR WAL. Você também não pode editar as configurações `hbase-site` para habilitar o Amazon EMR WAL em um cluster em execução.
- Você só pode habilitar o Amazon EMR WAL em clusters que usam o Amazon S3 como diretório raiz.
- Antes da versão 7.5.0 do Amazon EMR, os registros no Amazon EMR WAL precisavam ter 4 MB ou menos. Mas com o Amazon EMR versão 7.5.0 e posterior, o tamanho máximo do registro no EMR WAL é configurável, usando a propriedade. `emr.wal.max.payload.size` O valor padrão é de 1 GB. O exemplo a seguir define o tamanho máximo do registro como 2 GB:

```
[
  {
    "Classification":"hbase-site",
    "Properties": {
       "emr.wal.max.payload.size": "2147483648"
    }
  }
]
```

- Você não pode ter vários clusters ativos no mesmo diretório HBase raiz no Amazon S3.
- Você não pode habilitar o Amazon EMR WAL em clusters de réplica de leitura.
- O WAL é replicado em zonas de disponibilidade dentro do serviço gerenciado.
- O WAL sobrevive ao cluster e permanece disponível para o próximo.
- Você não pode desabilitar o Amazon EMR WAL durante a inicialização ou quando o cluster estiver em operação (em estado de execução).
- Para obter informações sobre o WAL e os limites do Workspace, consulte Amazon EMR endpoints and quotas.

Disponibilidade regional do Amazon EMR WAL

O serviço Amazon EMR WAL está disponível no seguinte: Regiões da AWS

- `ap-northeast-1`: Ásia-Pacífico (Tóquio)
- `ap-northeast-2`: Ásia-Pacífico (Seul)
- `ap-southeast-1`: Ásia-Pacífico (Singapura)
- `ap-south-1`: Ásia-Pacífico (Mumbai)
- `ap-southeast-2`: Ásia-Pacífico (Sydney)
- `eu-central-1`: Europa (Frankfurt)
- `eu-north-1`: Europa (Estocolmo)
- `eu-west-1`: Europa (Irlanda)
- `sa-east-1`: América do Sul (São Paulo)
- `us-east-1`: Leste dos EUA (Norte da Virgínia)
- `us-east-2`: Leste dos EUA (Ohio)
- `us-west-2`: Oeste dos EUA (Oregon)

As seguintes regiões só estariam disponíveis com a versão 7.3.0 ou posterior do Amazon EMR:

- `ap-east-1`: Ásia-Pacífico (Hong Kong)
- `af-south-1`: África (Cidade do Cabo)
- `ca-central-1`: Canadá (Central)
- `eu-west-2`: Europa (Londres)

Referência da CLI do Amazon EMR WAL (EMRWAL)

A interface da linha de comando do EMRWAL (EMRWAL CLI) é uma ferramenta unificada para gerenciar o log de gravação antecipada (WAL) do Amazon EMR. A EMRWAL CLI é fornecida com clusters do EMR ao habilitar o WAL no momento em que cria um cluster. Para obter mais informações sobre como habilitar o WAL, consulte Logs de gravação antecipada (WAL) do Amazon EMR.

A EMRWAL CLI inclui os seguintes comandos:

Tópicos

- createWorkspace
- deleteWal
- deleteWorkspace

- listTagsForResource
- listWals
- listWorkspaces
- tagResource
- untagResource

`createWorkspace`

O comando `createWorkspace` cria um Workspace do Amazon EMR WAL.

Uso:

```
emrwal createWorkspace [-tags <tags>] [-e {endpoint}] [-r {Region}] -w {workspacename}
 [-h]
```

Exemplo:

```
emrwal createWorkspace -w examplews
```

`deleteWal`

O comando `deleteWals` exclui o Amazon EMR WAL que você especifica.

Uso:

```
emrwal deleteWal [-e {endpoint}] [-r {Region}] [-w {workspacename}] [-p <tablePrefix>]
 [-n <walName>] [-N <fullName>] [-R] [-m] [-h]
```

Exemplo:

```
emrwal deleteWal -w examplews -p hbasetable -n examplewal
```

`deleteWorkspace`

O comando `deleteWorkspace` exclui o Workspace do Amazon EMR WAL que você especifica.

Uso:

```
emrwal deleteWorkspace [-e {endpoint}] [-r {Region}] -w {workspacename} [-h]
```

Exemplo:

```
emrwal deleteWorkspace -w examplews
```

`listTagsForResource`

O comando `listTagsForResource` lista todas as tags de pares de valores-chave do Workspace do Amazon EMR WAL que você especifica.

Uso:

```
emrwal listTagsForResource -arn {resource-arn} [-e {endpoint}] [-r {Region}] [-h]
```

Exemplo:

```
emrwal listTagsForResource -arn arn:aws:emrwal::1234567891234:workspace/examplews
```

`listWals`

O `listWals` comando lista todo o Amazon EMR WALs no espaço de trabalho que você especificar.

Uso:

```
emrwal listWals [-nextToken {token-string}] [-pageSize {integer}] [-e {endpoint}] [-
r {Region}] [-w {workspacename}] [-p <tablePrefix>] [-M {integer}] [-h]
```

Exemplo:

```
emrwal listWals -w examplews
```

`listWorkspaces`

O comando `listWorkspaces` lista todos os Workspaces do Amazon EMR WAL que estão disponíveis para você.

Uso:

```
emrwal listWorkspaces [-nextToken {token-string}] [-pageSize {integer}] [-e {endpoint}]
 [-r {Region}] [-M {integer}] [-h]
```

Exemplo:

```
emrwal listWorkspaces
```

`tagResource`

O comando `tagResource` atribui uma ou mais tags de pares de valores-chave ao Workspace do Amazon EMR WAL que você especifica.

Uso:

```
emrwal tagResource -arn {resource-arn} -tags <tags> [-e {endpoint}] [-r {Region}] [-h]
```

Exemplo:

```
emrwal tagResource -arn arn:aws:emrwal::1234567891234:workspace/examplews -
tags tag_key=tag_value
```

`untagResource`

O comando `untagResource` cancela a atribuição de uma ou mais tags de pares de valores-chave ao Workspace do Amazon EMR WAL que você especifica.

Uso:

```
emrwal untagResource -arn {resource-arn} -tagKeys <tagKeys> [-e {endpoint}] [-
r {Region}] [-h]
```

Exemplo:

```
emrwal untagResource -arn arn:aws:emrwal::1234567891234:workspace/examplews -
tagKeys tag_key
```

Usando a HBase concha

Depois de criar um HBase cluster, a próxima etapa é conectar-se para que você HBase possa começar a ler e gravar dados (as gravações de dados não são suportadas em um cluster de réplica de leitura). Você pode usar o HBase shell para testar comandos.

Para abrir a HBase concha

1. Use SSH para se conectar ao servidor principal no HBase cluster. Para obter informações sobre como se conectar ao nó primário do Amazon EMR usando SSH, consulte Connect to the primary node using SSH no Guia de gerenciamento do Amazon EMR.
2. Executar `hbase shell`. O HBase shell é aberto com um prompt semelhante ao seguinte.

```
hbase(main):001:0>
```

Você pode emitir comandos HBase shell a partir do prompt. Para obter mais informações sobre os comandos do shell e como chamá-los, digite help no HBase prompt e pressione Enter.

Criar uma tabela

O comando a seguir cria uma tabela chamada “t1”, que tem uma família de coluna exclusiva chamada de “f1”.

```
hbase(main):001:0>create 't1', 'f1'
```

Inserir um valor

O comando a seguir insere o valor “v1” para a linha “r1” na tabela “t1” e na coluna “f1”.

```
hbase(main):001:0>put 't1', 'r1', 'f1:col1', 'v1'
```

Obter um valor

O comando a seguir obtém os valores para a linha “r1” na tabela “t1”.

```
hbase(main):001:0>get 't1', 'r1'
```

Excluir uma tabela

O comando a seguir descarta e exclui a tabela “t1”.

```
hbase(main):001:0>drop 'ns1:t1',false
```

O valor booleano indica se você deseja ou não arquivar a tabela, então você pode defini-la como `true` se quiser salvá-la. Também é possível executar `drop 'ns1:t1'` sem booleano para arquivar a tabela.

Acesse HBase tabelas com o Hive

HBase e Apache Hive estão totalmente integrados, permitindo a execução massiva de cargas de trabalho de processamento paralelo diretamente nos dados armazenados em. HBase Para usar o Hive com HBase, geralmente você pode iniciá-los no mesmo cluster. No entanto, você pode iniciar o Hive e HBase em clusters separados. Executar HBase o Hive separadamente em clusters diferentes pode melhorar o desempenho, pois isso permite que cada aplicativo use os recursos do cluster com mais eficiência.

Os procedimentos a seguir mostram como se conectar HBase em um cluster usando o Hive.

> **Note**
>
> Você só pode conectar um cluster do Hive a um único HBase cluster.

Para conectar o Hive ao HBase

1. Crie clusters separados com o Hive e HBase instalado ou crie um único cluster com ambos HBase e o Hive instalados.
2. Se você estiver usando clusters separados, modifique seus grupos de segurança para que HBase as portas do Hive estejam abertas entre esses dois nós primários.
3. Use o SSH para se conectar ao nó central do cluster com o Hive instalado. Para obter mais informações, consulte Connect to the primary node using SSH no Guia de gerenciamento do Amazon EMR.
4. Execute o shell do Hive com o comando a seguir.

   ```
   hive
   ```

5. (Opcional) Não é necessário fazer isso se HBase o Hive estiver localizado no mesmo cluster. Conecte o HBase cliente em seu cluster Hive ao HBase cluster que contém seus dados. No exemplo a seguir, *public-DNS-name* é substituído pelo nome DNS público do nó primário do HBase cluster, por exemplo:`ec2-50-19-76-67.compute-1.amazonaws.com`.

```
set hbase.zookeeper.quorum=public-DNS-name;
```

6. Continue executando as consultas do Hive em seus HBase dados conforme desejado ou veja o próximo procedimento.

Para acessar HBase dados do Hive

- Depois que a conexão entre o Hive e os HBase clusters for estabelecida (conforme mostrado no procedimento anterior), você poderá acessar os dados armazenados no HBase cluster criando uma tabela externa no Hive.

  O exemplo a seguir, quando executado a partir do prompt do Hive no nó primário, cria uma tabela externa que faz referência aos dados armazenados em uma HBase tabela chamada`inputTable`. Em seguida, você pode fazer referência `inputTable` nas instruções do Hive para consultar e modificar os dados armazenados no HBase cluster.

  ```
  set hbase.zookeeper.quorum=ec2-107-21-163-157.compute-1.amazonaws.com;

  create external table inputTable (key string, value string)
       stored by 'org.apache.hadoop.hive.hbase.HBaseStorageHandler'
        with serdeproperties ("hbase.columns.mapping" = ":key,f1:col1")
        tblproperties ("hbase.table.name" = "t1");

  select count(key) from inputTable ;
  ```

Para um caso de uso mais avançado e um exemplo de combinação com o Hive, consulte a postagem do blog sobre AWS Big Data, Combine NoSQL HBase e análises paralelas massivas usando o Apache e o HBase Apache Hive no Amazon EMR.

Usando HBase instantâneos

HBase usa uma funcionalidade de captura instantânea integrada para criar backups leves de tabelas. Em clusters EMR, esses backups podem ser exportados para o Amazon S3 usando o EMRFS. Você pode criar um instantâneo no nó primário usando o HBase shell. Este tópico mostra

como executar esses comandos interativamente com o shell ou por meio de uma etapa usando `command-runner.jar` o AWS CLI ou AWS SDK para Java. Para obter mais informações sobre outros tipos de HBase backups, consulte HBase backup na HBase documentação.

Criar um snapshot usando uma tabela

```
hbase snapshot create -n snapshotName -t tableName
```

Usando o command-runner.jar a partir do AWS CLI:

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps Name="HBase Shell Step",Jar="command-runner.jar",\
Args=[ "hbase", "snapshot", "create","-n","snapshotName","-t","tableName"]
```

AWS SDK para Java

```
HadoopJarStepConfig hbaseSnapshotConf = new HadoopJarStepConfig()
  .withJar("command-runner.jar")
  .withArgs("hbase","snapshot","create","-n","snapshotName","-t","tableName");
```

> **Note**
>
> Se o nome do snapshot não for exclusivo, a operação de criação falhará com um código de retorno `-1` ou `255`, mas talvez você não veja uma mensagem de erro afirmando o que deu errado. Para usar o mesmo nome de snapshot, exclua-o e recrie-lo em seguida.

Excluir um snapshot

```
hbase shell
>> delete_snapshot 'snapshotName'
```

Exibir informações do snapshot

```
hbase snapshot info -snapshot snapshotName
```

Exportar um snapshot para o Amazon S3

> **Important**
>
> Se você não especificar um `-mappers` valor ao exportar um instantâneo, HBase usa um cálculo arbitrário para determinar o número de mapeadores. Esse valor pode ser muito grande, dependendo do tamanho da tabela, o que afeta negativamente os trabalhos em execução durante a exportação. Por esse motivo, recomendamos que você especifique o parâmetro `-mappers`, o parâmetro `-bandwidth` (que indica o consumo de largura de banda em megabytes por segundo) ou ambos para limitar os recursos de cluster usados pela operação de exportação. Como alternativa, você pode executar a operação de exportação de snapshot durante um período de baixa utilização.

```
hbase snapshot export -snapshot snapshotName \
-copy-to s3://amzn-s3-demo-bucket/folder -mappers 2
```

Usando `command-runner.jar` a partir do AWS CLI:

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps Name="HBase Shell Step",Jar="command-runner.jar",\
Args=[ "hbase", "snapshot", "export","-snapshot","snapshotName","-copy-to","s3://amzn-s3-demo-bucket/folder","-mappers","2","-bandwidth","50"]
```

AWS SDK para Java:

```
HadoopJarStepConfig hbaseImportSnapshotConf = new HadoopJarStepConfig()
  .withJar("command-runner.jar")
  .withArgs("hbase","snapshot","export",
      "-snapshot","snapshotName","-copy-to",
      "s3://bucketName/folder",
      "-mappers","2","-bandwidth","50");
```

Importar um snapshot do Amazon S3

Embora seja uma importação, a HBase opção usada aqui ainda é`export`.

```
sudo -u hbase hbase snapshot export \
-D hbase.rootdir=s3://amzn-s3-demo-bucket/folder \
```

```
-snapshot snapshotName \
-copy-to hdfs://masterPublicDNSName:8020/user/hbase \
-mappers 2
```

Usando `command-runner.jar` a partir do AWS CLI:

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps Name="HBase Shell Step",Jar="command-runner.jar", \
Args=["sudo","-u","hbase","hbase snapshot export","-snapshot","snapshotName", \
"-D","hbase.rootdir=s3://amzn-s3-demo-bucket/folder", \
"-copy-to","hdfs://masterPublicDNSName:8020/user/hbase","-mappers","2","-chmod","700"]
```

AWS SDK para Java:

```
HadoopJarStepConfig hbaseImportSnapshotConf = new HadoopJarStepConfig()
  .withJar("command-runner.jar")
  .withArgs("sudo","-u","hbase","hbase","snapshot","export", "-D","hbase.rootdir=s3://
path/to/snapshot",
      "-snapshot","snapshotName","-copy-to",
      "hdfs://masterPublicDNSName:8020/user/hbase",
      "-mappers","2","-chuser","hbase");
```

Restaurar uma tabela a partir de instantâneos dentro do shell HBase

```
hbase shell
>> disable tableName
>> restore_snapshot snapshotName
>> enable tableName
```

HBase atualmente não oferece suporte a todos os comandos de snapshot encontrados no HBase shell. Por exemplo, não há opção de HBase linha de comando para restaurar um snapshot, então você deve restaurá-lo em um shell. Isso significa que `command-runner.jar` deve executar um comando Bash.

> **Note**
>
> Como o comando usado aqui é `echo`, é possível que seu comando shell ainda falhe, mesmo que o comando executado pelo Amazon EMR retorne um código de saída `0`. Verifique os logs de etapas se você optar por executar um comando shell como um etapa.

```
echo 'disable tableName; \
restore_snapshot snapshotName; \
enable tableName' | hbase shell
```

Veja a seguir a etapa usando a AWS CLI. Primeiro, crie o seguinte arquivo `snapshot.json`:

```
[
  {
    "Name": "restore",
    "Args": ["bash", "-c", "echo $'disable \"tableName\"; restore_snapshot
 \"snapshotName\"; enable \"tableName\"' | hbase shell"],
    "Jar": "command-runner.jar",
    "ActionOnFailure": "CONTINUE",
    "Type": "CUSTOM_JAR"
  }
]
```

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps file://./snapshot.json
```

AWS SDK para Java:

```
HadoopJarStepConfig hbaseRestoreSnapshotConf = new HadoopJarStepConfig()
  .withJar("command-runner.jar")
  .withArgs("bash","-c","echo $'disable \"tableName\"; restore_snapshot \"snapshotName
\"; enable \"snapshotName\"' | hbase shell");
```

Configure HBase

Embora as HBase configurações padrão devam funcionar para a maioria dos aplicativos, você pode modificá-las. HBase Para fazer isso, use as propriedades das classificações de HBase configuração. Para obter mais informações, consulte Configurar aplicações.

O exemplo a seguir cria um cluster com um diretório HBase raiz alternativo baseado em um arquivo de configuração,`myConfig.json`, armazenado no Amazon S3.

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr create-cluster --release-label emr-7.9.0 --applications Name=HBase \
--instance-type m5.xlarge --instance-count 3 --configurations https://s3.amazonaws.com/
amzn-s3-demo-bucket/myfolder/myConfig.json
```

O arquivo `myConfig.json` especifica a propriedade `hbase.rootdir` para a classificação de configuração `hbase-site`, conforme mostrado no exemplo a seguir. *`ip-XXX-XX-XX-XXX.ec2.internal`*Substitua pelo nome de host DNS interno do nó primário do cluster.

```
[
  {
    "Classification":"hbase-site",
    "Properties": {
       "hbase.rootdir": "hdfs://ip-XXX-XX-XX-XXX.ec2.internal:8020/user/
myCustomHBaseDir"
    }
  }
]
```

> **Note**
>
> Com as versões 5.21.0 e posteriores do Amazon EMR, você pode substituir as configurações de cluster e especificar classificações de configuração adicionais para cada grupo de instâncias em um cluster em execução. Você faz isso usando o console do Amazon EMR, o AWS Command Line Interface (AWS CLI) ou o AWS SDK. Para obter mais informações, consulte Supplying a Configuration for an Instance Group in a Running Cluster.

Alterações na alocação de memória do YARN

HBase não está sendo executado como um aplicativo YARN, portanto, é necessário recalcular a memória alocada para o YARN e seus aplicativos, o que resulta em uma redução na memória

geral disponível para o YARN se estiver instalado. HBase Você deve levar isso em consideração ao planejar a co-localização de aplicativos YARN e HBase nos mesmos clusters. Os tipos de instância com menos de 64 GB de memória têm metade da memória disponível paraNodeManager, que é então alocada para o. HBase RegionServer Por exemplo, tipos com memória maior que 64 GB, a HBase RegionServer memória é limitada a 32 GB. Como regra geral, a memória de configuração do YARN é um múltiplo da memória de tarefas do MapReduce redutor.

As tabelas em Valores padrão para definições de configuração de tarefa mostram as alterações nas configurações do YARN com base na memória necessária para HBase.

HBase números de porta

Alguns números de porta escolhidos HBase são diferentes do padrão. A seguir estão as interfaces e portas para o HBase Amazon EMR.

HBase portas

Interface	Porta	Protocolo
HMaster	16000	TCP
HMaster UI	16010	HTTP
RegionServer	16020	TCP
RegionServer Informações	16030	HTTP
Servidor REST	8070	HTTP
INTERFACE DO USUÁRIO REST	8085	HTTP
Servidor Thrift	9090	TCP
Interface do usuário do servidor Thrift	9095	HTTP

> **Important**
>
> O `kms-http-port` é 9700 e o `kms-admin-port` é 9701 no Amazon EMR 4.6.0 e versões posteriores.

HBase configurações do site para otimizar

Você pode definir qualquer uma ou todas as configurações do HBase site para otimizar o HBase cluster para a carga de trabalho do seu aplicativo. Recomendamos as seguintes configurações como ponto de partida na sua investigação.

zookeeper.session.timeout

O tempo limite padrão é de 40 segundos (40.000 ms). Se um servidor de regiões travar, este será o tempo necessário para o servidor mestre notar a ausência do servidor de regiões e iniciar a recuperação. Para ajudar o servidor mestre a se recuperar com mais rapidez, você pode reduzir esse valor para um período mais curto. O exemplo a seguir usa 30 segundos ou 30.000 ms:

```
[
  {
    "Classification":"hbase-site",
    "Properties": {
       "zookeeper.session.timeout": "30000"
    }
  }
]
```

hbase.regionserver.handler.count

Define o número de threads que o servidor de regiões mantém abertos para atender às solicitações de tabelas. O padrão de 10 é baixo, a fim de impedir que os usuários eliminem seus servidores de regiões ao usarem buffers de gravação grandes com um alto número de clientes simultâneos. A regra geral é manter esse número baixo quando a carga útil por solicitação se aproxima da faixa de MB (grandes entradas, digitalizações usando um cache grande) e alto quando a carga útil é pequena (obtenções, pequenas entradas, ICVs exclusões). O exemplo a seguir aumenta o número de threads abertos para 30:

```
[
```

```
  {
    "Classification":"hbase-site",
    "Properties": {
       "hbase.regionserver.handler.count": "30"
    }
  }
]
```

hbase.hregion.max.filesize

Esse parâmetro determina o tamanho, em bytes, das regiões individuais. Por padrão, ele é definido como `1073741824`. Se você estiver gravando muitos dados em seu HBase cluster e isso estiver causando divisões frequentes, você pode aumentar esse tamanho para aumentar as regiões individuais. Isso reduz as divisões, mas exige mais tempo para fazer o balanceamento de carga de regiões de um servidor para outro.

```
[
  {
    "Classification":"hbase-site",
    "Properties": {
       "hbase.hregion.max.filesize": "1073741824"
    }
  }
]
```

hbase.hregion.memstore.flush.size

Esse parâmetro determina o tamanho máximo de memstore, em bytes, antes que ele seja liberado no disco. O padrão é `134217728`. Se a sua workload é formada por curtos disparos contínuos de operações de gravação, convém aumentar esse limite para que todas as gravações permaneçam na memória durante o disparo contínuo e sejam liberadas no disco mais tarde. Isso pode aumentar o desempenho durante disparos contínuos.

```
[
  {
    "Classification":"hbase-site",
    "Properties": {
       "hbase.hregion.memstore.flush.size": "134217728"
    }
  }
]
```

Considerações sobre a performance

Habilitando um coletor de lixo (ZGC) para HBase

Com o Amazon EMR versão 7.10.0 e posterior, os usuários podem configurar facilmente as configurações de coleta de lixo (GC) para seu cluster. HBase Recomendamos ativar o Z Garbage Collector (ZGC) para obter tempos de pausa GC de baixa latência e menos de um milissegundo.

Aqui está uma configuração para habilitar o ZGC para HBase RegionServer (s):

```
[
    {
        "Classification": "hbase-env",
        "Properties": {},
        "Configurations": [
            {
                "Classification": "export",
                "Properties": {
                    "JAVA_HOME": "/usr/lib/jvm/jre-21",
                    "HBASE_REGIONSERVER_GC_OPTS": "\"-XX:+UseZGC -XX:+ZGenerational\""
                }
            }
        ]
    }
]
```

> **Note**
>
> A variável de ambiente JAVA_HOME deve ser definida ao usar o ZGC geracional (recomendado) porque foi introduzida no JDK 21. Se você quiser usar o modo não geracional (sem **`-XX:+ZGenerational`**) do ZGC, não há necessidade de definir o JAVA_HOME. No JDK 24, o modo não geracional do ZGC foi removido.

Ajuste do ZGC

1. Ativar pilha fixa da JVM

   O Z Garbage Collector (ZGC) tem um desempenho mais eficiente quando configurado com memória de pilha fixa, eliminando a sobrecarga de devolução da memória ao sistema operacional. Para configurar a memória de pilha fixa para seu HBase cluster, use a seguinte configuração:

```
[
    {
        "Classification": "hbase",
        "Properties": {
            "hbase.regionserver.fixed.heap.enabled": "true"
        }
    }
]
```

2. Ativar pré-toque

   A ativação do pré-toque melhora o desempenho da coleta de lixo (GC), reduzindo a latência inicial e fornecendo um desempenho mais previsível. Para habilitar o pré-toque para seu HBase cluster, use a seguinte configuração:

```
[
    {
        "Classification": "hbase-env",
        "Properties": {},
        "Configurations": [
            {
                "Classification": "export",
                "Properties": {
                    "JAVA_HOME": "/usr/lib/jvm/jre-21",
                    "HBASE_REGIONSERVER_GC_OPTS": "\"-XX:+UseZGC -XX:+ZGenerational -
XX:+AlwaysPreTouch\""
                }
            }
        ]
    }
]
```

Visualize a interface HBase do usuário

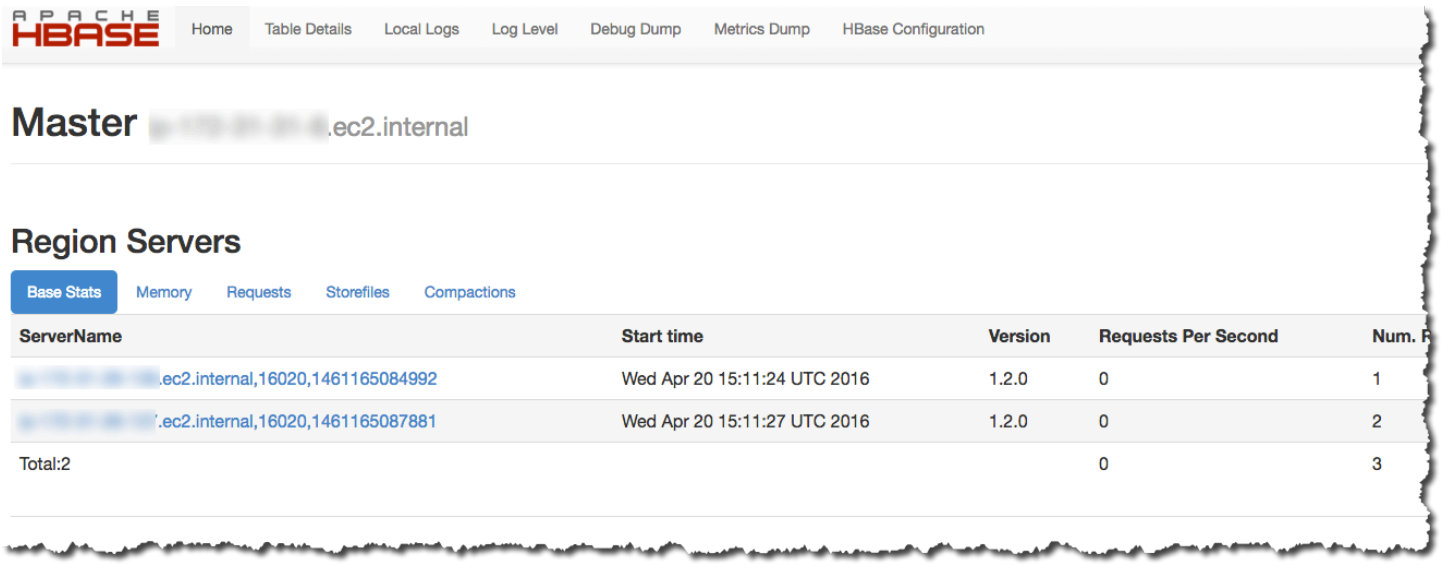
> **Note**
>
> A interface HBase do usuário usa conexões HTTP inseguras por padrão. Para habilitar o HTTP seguro (HTTPS), defina a `hbase.ssl.enabled` propriedade da `hbase-site` classificação `true` em sua HBase configuração. Para obter mais informações sobre o uso de

HTTP seguro (HTTPS) para a interface do usuário HBase da web, consulte o Guia de HBase referência do Apache.

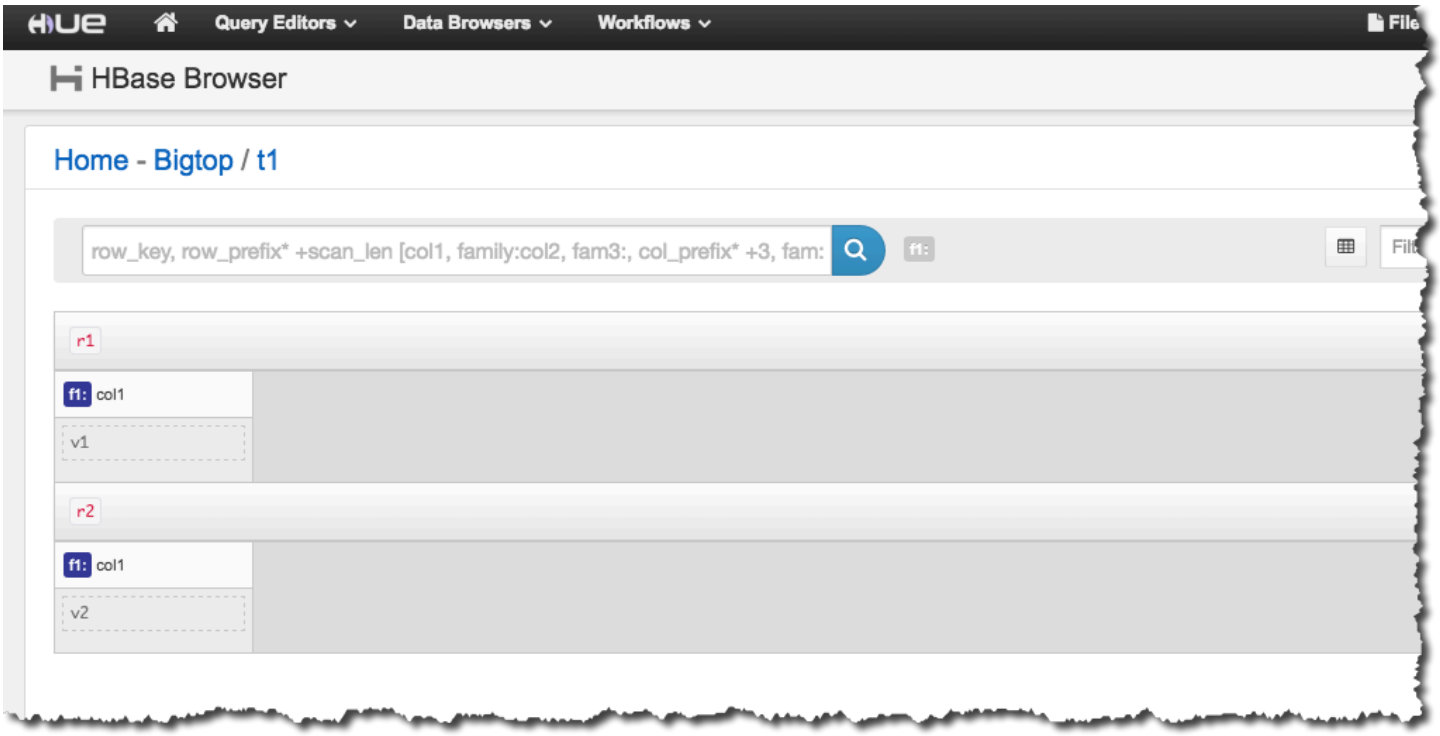
HBase fornece uma interface de usuário baseada na web que você pode usar para monitorar seu HBase cluster. Quando você executa HBase no Amazon EMR, a interface web é executada no nó primário e pode ser visualizada usando o encaminhamento de portas, também conhecido como criação de um túnel SSH.

Para visualizar a interface HBase do usuário

1. Use o SSH para formar um túnel até o nó primário e criar uma conexão segura. Para obter mais informações, consulte Opção 2, parte 1: configurar um túnel SSH para o nó primário usando o encaminhamento dinâmico de portas no Guia de gerenciamento do Amazon EMR.
2. Instale um navegador da Web com uma ferramenta de proxy, como o FoxyProxy plug-in do Firefox, para criar um proxy SOCKS para AWS domínios. Para obter mais informações, consulte Opção 2, parte 2: Definir configurações de proxy para visualizar sites hospedados no nó primário no Guia de gerenciamento do Amazon EMR.
3. Com o proxy definido e a conexão SSH aberta, você pode visualizar a HBase interface do usuário abrindo uma janela do navegador com http: //:16010/master-status***master-public-dns-name***, onde está *master-public-dns-name* o endereço DNS público do nó primário do cluster.



Você também pode ver HBase em Hue. Por exemplo, o seguinte mostra a tabela, `t1`, criada em Usando a HBase concha:



Para obter mais informações sobre o Hue, consulte Hue.

Exibir arquivos HBase de log

Como parte de sua operação, HBase grava arquivos de log com detalhes sobre configurações, ações do daemon e exceções. Esses arquivos de log podem ser úteis para depurar problemas e HBase monitorar o desempenho.

Se você configurar o cluster para manter a persistência de arquivos de log para o Amazon S3, deve saber que os logs são gravados no Amazon S3 cada cinco minutos e, portanto, pode haver um pequeno atraso antes que os arquivos de log mais recentes estejam disponíveis.

Para visualizar HBase registros no nó primário

- Você pode visualizar os HBase registros atuais usando SSH para se conectar ao nó primário e navegar até o `/var/log/hbase` diretório. Esses logs não estão disponíveis depois que o cluster é encerrado, a menos que você habilite o registro em log no Amazon S3 quando o cluster for iniciado.

Para visualizar HBase registros no Amazon S3

- Para acessar HBase logs e outros logs de cluster no Amazon S3 e disponibilizá-los após o encerramento do cluster, especifique um bucket do Amazon S3 para receber esses registros ao criar o cluster. Isso é feito usando a opção `--log-uri`. Para obter mais informações sobre como habilitar o registro em log para o seu cluster, consulte Configure logging and debugging (optional) no Guia de gerenciamento do Amazon EMR.

Monitor HBase com Ganglia

> **Note**
>
> A última versão do Amazon EMR a incluir o Ganglia foi o Amazon EMR 6.15.0. Para monitorar seu cluster, versões superiores à 6.15.0 incluem o agente Amazon CloudWatch .

O projeto de código-fonte aberto Ganglia é um sistema distribuído e escalável, projetado para monitorar clusters e grades e, ao mesmo tempo, minimizar o impacto sobre o seu desempenho. Quando você habilita o Ganglia no seu cluster, pode gerar relatórios e visualizar o desempenho do cluster como um todo, bem como inspecionar o desempenho de instâncias de nós individuais. Para obter mais informações sobre o projeto de código-fonte aberto Ganglia, consulte http://ganglia.info/. Para obter mais informações sobre como usar o Ganglia com clusters do Amazon EMR, consulte Ganglia.

Depois que o cluster é iniciado com o Ganglia configurado, você pode acessar os gráficos e relatórios do Ganglia usando a interface gráfica em execução no nó primário.

O Ganglia armazena arquivos de log no nó primário no diretório `/mnt/var/lib/ganglia/rrds/`. Versões anteriores do Amazon EMR podem armazenar arquivos de log no `/var/log/ganglia/rrds/` diretório.

Para configurar um cluster para o Ganglia e usar o HBase AWS CLI

- Use um comando `create-cluster` semelhante ao seguinte:

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.9.0 \
--applications Name=HBase Name=Ganglia --use-default-roles \
--ec2-attributes KeyName=myKey --instance-type m5.xlarge \
```

```
--instance-count 3
```

> **Note**
>
> Se a função de serviço padrão do Amazon EMR e o perfil de EC2 instância da Amazon não existirem, ocorrerá um erro. Use o comando `aws emr create-default-roles` para criá-los e, em seguida, tente novamente.

Para obter mais informações, consulte os comandos do Amazon EMR na AWS CLI.

Para visualizar HBase métricas na interface web do Ganglia

1. Use o SSH para formar um túnel até o nó primário e criar uma conexão segura. Para obter mais informações, consulte Opção 2, parte 1: configurar um túnel SSH para o nó primário usando o encaminhamento dinâmico de portas no Guia de gerenciamento do Amazon EMR.
2. Instale um navegador da Web com uma ferramenta de proxy, como o FoxyProxy plug-in do Firefox, para criar um proxy SOCKS para AWS domínios. Para obter mais informações, consulte Opção 2, parte 2: Definir configurações de proxy para visualizar sites hospedados no nó primário no Guia de gerenciamento do Amazon EMR.
3. Com o proxy definido e a conexão SSH aberta, você pode visualizar as métricas do Ganglia abrindo uma janela do navegador com http://*master-public-dns-name*/ganglia/, onde *master-public-dns-name* está o endereço DNS público do servidor mestre no cluster. HBase

Para exibir arquivos de log do Ganglia no nó central

- Se o cluster ainda estiver em execução, você poderá acessar os arquivos de log usando o SSH para se conectar ao nó primário e navegar até o diretório `/mnt/var/lib/ganglia/rrds/`. No EMR 3.x, navegue até o diretório `/var/log/ganglia/rrds`. Para obter mais informações, consulte Connect to the primary node using SSH no Guia de gerenciamento do Amazon EMR.

Visualizar arquivos de log do Ganglia no Amazon S3

- Os arquivos de log do Ganglia não são gravados automaticamente no Amazon S3 mesmo se você habilitar o registro em log para seu cluster. Para visualizar os arquivos de log do Ganglia no

Amazon S3, envie manualmente os logs do `/mnt/var/lib/ganglia/rrds/` para o bucket do S3.

Monitorando o EMR com HBase a Amazon CloudWatch

A partir do EMR 7.0, o Amazon EMR fornece o CloudWatch Amazon Agent para enviar métricas para o Prometheus CloudWatch ou para o Prometheus, substituindo o sistema de monitoramento Ganglia. Você pode encontrar mais detalhes no Amazon CloudWatch Agent Configuration Guide.

O agente CloudWatch Amazon do EMR 7.0 forneceu integração básica e sua configuração exigiu esforço manual, como baixar componentes e usar ações de bootstrap. No entanto, a partir do EMR 7.1, o processo é simplificado. O uso do Amazon CloudWatch Agent no EMR 7.1 facilita o monitoramento de métricas específicas do sistema e do aplicativo, incluindo aquelas para. HBase Ao usar a API de configuração do EMR, você pode configurar e personalizar rapidamente o processo de coleta de métricas e escolher para onde enviar os dados, seja para a Amazon ou para a Prometheus. CloudWatch Essa flexibilidade ajuda você a acompanhar de perto seus HBase clusters, garantindo que eles funcionem sem problemas e com eficiência.

Usando a API de configuração do EMR para lidar com toda a configuração, tornando o processo muito mais fácil. O CloudWatch agente no EMR 7.1 oferece suporte a três tipos principais de métricas:

- Métricas do sistema — incluem indicadores-chave do desempenho do sistema, como: uso de CPU, uso de disco, uso de memória, E/S de rede, processos e uso de swap.
- Métricas do Hadoop Daemon — Essas métricas estão relacionadas aos vários componentes do Hadoop, incluindo: DataNode métricas, métricas, métricas do YARN e métricas do NodeManager YARN. NameNode ResourceManager
- HBase Métricas — Essas métricas fornecem informações sobre o desempenho HBase do: métricas HBase principais, métricas do servidor HBase regional, métricas do servidor HBase REST e métricas do HBase Thrift Server.

Using the AWS CLI

Todas as métricas do Hadoop HBase são baseadas em JMX, o que significa que elas usam extensões de gerenciamento Java para fornecer informações detalhadas. Veja como você pode configurar o Amazon CloudWatch Agent para monitorar HBase:

- Consulte os pré-requisitos antes de criar um cluster com o Amazon Agent. CloudWatch Use um `create-cluster` comando semelhante ao exemplo que aparece após essa lista.
- Consulte as configurações compatíveis com um HBase cluster.
- Consulte o exemplo a seguir para ver exemplos de configurações para configurar o HBase monitoramento. Consulte as configurações de exemplo para a `--configuration` entrada.

```
aws emr create-cluster --name "HBase cluster with CloudWatch agent" \
--release-label emr-7.1.0 \
--applications Name=HBase Name=AmazonCloudWatchAgent \
--ec2-attributes KeyName=myKey --instance-type m7g.2xlarge \
--configurations file://./configurations.json \
--instance-count 3 --use-default-roles
```

Para obter mais informações sobre configurações de amostra JSON para métricas, consulte Configurar métricas.

Using the console

Para criar um cluster com o CloudWatch agente da Amazon a partir do console, execute estas etapas:

Crie um cluster com o CloudWatch agente a partir do console

1. Faça login no e abra AWS Management Console o console do Amazon EMR em https://console.aws.amazon.com /emr.
2. Selecione Criar cluster.
3. Em Nome e aplicações, escolha uma versão 7.0.0 ou superior do Amazon EMR.
4. Em Pacote de aplicativos, selecione HBase os aplicativos adicionais que você deseja instalar em seu cluster e inclua o CloudWatch agente em suas seleções.
5. Expanda as configurações de software. Você poderá inserir a configuração diretamente usando JSON ou uma sintaxe abreviada demonstrada no texto sombreado no console. Caso contrário, você pode fornecer um URI do Amazon S3 para um arquivo com um objeto JSON `Configurations`. Para obter mais informações sobre configurações de amostra JSON para métricas, consulte Configurar métricas.
6. Prossiga com a criação do cluster para atender às necessidades do seu caso de uso.

Configurar métricas

Para monitorar o HBase Master, você pode configurar o Amazon CloudWatch Agent para coletar métricas específicas.

1. Configurando métricas HBase principais — Para monitorar o HBase Master, você pode configurar o Amazon CloudWatch Agent para coletar métricas específicas. Aqui está um exemplo de configuração para monitorar as atividades do Master Assignment Manager:

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-hbase-master-metrics",
        "Properties": {
          "Hadoop:service=HBase,name=Master,sub=AssignmentManager":
 "AssignFailedCount,AssignSubmittedCount",
          "otel.metric.export.interval": "30000"
        },
        "Configurations": []
      }
    ]
  }
]
```

   Nesta configuração:

   - Especificamos o MBean (`Hadoop:service=HBase,name=Master,sub=AssignmentManager`) para coletar métricas como `AssignFailedCount` `AssignSubmittedCount` e.
   - Definimos o intervalo para coletar essas métricas a cada 30 segundos (30.000 milissegundos).
2. Configurando métricas do servidor de HBase região — Para monitorar servidores de HBase região, configure o CloudWatch agente da seguinte forma:

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
```

```
    "Configurations": [
      {
        "Classification": "emr-hbase-region-server-metrics",
        "Properties": {
          "Hadoop:service=HBase,name=RegionServer,sub=IPC":
 "numActiveHandler,numActivePriorityHandler",
          "otel.metric.export.interval": "30000"
        },
        "Configurations": []
      }
    ]
  }
]
```

Essa configuração:

- Monitora manipuladores ativos no Servidor de Região (`numActiveHandler`,`numActivePriorityHandler`).
- Usa um intervalo de 30 segundos para coleta de métricas.

3. Configurando métricas do servidor HBase REST — Para monitorar a interface HBase REST, você pode usar a seguinte configuração:

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-hbase-rest-server-metrics",
        "Properties": {
          "Hadoop:service=HBase,name=REST": "successfulPut,successfulScanCount",
          "otel.metric.export.interval": "30000"
        },
        "Configurations": []
      }
    ]
  }
]
```

Neste exemplo, o CloudWatch agente coleta métricas sobre operações PUT bem-sucedidas e contagens de escaneamento a cada 30 segundos.

4. Configurando métricas do HBase Thrift Server — Para monitorar o HBase Thrift Server, você pode configurar métricas com uma configuração como a seguinte:

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-hbase-thrift-server-metrics",
        "Properties": {
          "Hadoop:service=HBase,name=Thrift,sub=ThriftOne":
 "BatchGet_max,BatchGet_mean",
          "otel.metric.export.interval": "30000"
        },
        "Configurations": []
      }
    ]
  }
]
```

Essa configuração rastreia os tempos máximos e médios das operações GET em lote no servidor Thrift.

Usando o destino de métricas

Com o EMR 7.1, você tem a opção de enviar seus dados de métricas para a Amazon ou o CloudWatch Amazon Managed Service for Prometheus. Essa opção permite que você se integre perfeitamente a diferentes ferramentas de monitoramento, com base em suas necessidades.

Enviando métricas para a Amazon CloudWatch

Para enviar métricas para CloudWatch, use esta configuração:

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {
      "metrics_destination": "cloudwatch"
    },
    "Configurations": []
```

```
  }
]
```

Envio de métricas para o Amazon Managed Service for Prometheus

Se você preferir usar o Prometheus, defina o destino e forneça a URL do endpoint:

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {
      "metrics_destination": "prometheus",
      "prometheus_endpoint": "https://aps-workspaces.region.amazonaws.com/
workspaces/workspace_id/api/v1/remote_write"
    },
    "Configurations": []
  }
]
```

`region`Substitua pela sua AWS região e pelo `workspace_id` ID do espaço de trabalho Prometheus. Essa configuração direciona suas HBase métricas para o Prometheus usando o endpoint especificado.

Com a configuração acima, você pode ver as métricas abaixo na guia Monitoramento.

Migrando de versões anteriores HBase

Para migrar dados de uma HBase versão anterior, consulte Atualização, número da HBase versão e compatibilidade no Guia de referência do Apache HBase . Talvez seja necessário prestar atenção especial aos requisitos de atualização das versões anteriores à 1.0 do. HBase

Migração para o Amazon EMR versão 7.4.0 ou posterior

> **Note**
>
> Siga estas diretrizes se você estiver migrando de uma versão do EMR anterior à 7.4.0 para uma versão superior à 7.3.0.

Se você estiver executando atualmente uma versão do EMR com o recurso de rastreamento de arquivos da loja da Amazon ativado, incluído nas versões 6.2.0 a 7.3.0, e quiser fazer o upgrade para uma versão com o OSS Store File Tracking, disponível nas versões do EMR posteriores à 7.3.0, siga estas etapas:

1. No cluster existente:
    1. Desative a `hbase:storefile` tabela.
    2. Descarte a tabela `hbase:storefile`.
    3. Lavar`hbase:meta`.
    4. Aguarde até que os metadados sejam atualizados.
2. No novo cluster:
    1. Defina o mesmo diretório do Amazon S3 como o diretório raiz.
    2. Inicie o cluster com a `DefaultStoreFileTracker` implementação:

```
{
  "Classification": "hbase-site",
  "Properties": {
    hbase.store.file-tracker.impl:
 "org.apache.hadoop.hbase.regionserver.storefiletracker.DefaultStoreFileTracker"
}
```

    3. No nível da tabela ou do grupo de colunas, use os seguintes comandos para alterar o rastreador de arquivos armazenados:
        a. Altere o Store File Tracker da tabela ou da família de colunas da tabela:

```
hbase> change_sft 't1','FILE'
hbase> change_sft 't2','cf1','FILE'
```

        b. Altere todo o Store File Tracker da tabela que corresponda à expressão regular fornecida (regex):

```
hbase> change_sft_all 't.*','FILE'
hbase> change_sft_all 'ns:.*','FILE'
hbase> change_sft_all 'ns:t.*','FILE'
```

HBase histórico de lançamentos

A tabela a seguir lista a versão HBase incluída em cada versão de lançamento do Amazon EMR, junto com os componentes instalados com o aplicativo. Para obter as versões dos componentes em cada versão, consulte a seção Versão de componentes da sua versão em Versões 7.x do Amazon EMR, Versões de lançamento 6.x do Amazon EMR ou Versões de lançamento 5.x do Amazon EMR.

HBase informações sobre a versão

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-7.9.0	2.6.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server
emr-7.8.0	2.6.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-ma

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
		pred, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server, hbase-hma ster, hbase-client, hbase-reg ion-server, hbase-rest-server, hbase-thrift-server, hbase-ope rator-tools, zookeeper-client, zookeeper-server
emr-7.7.0	2.6.1	emrfs, emr-ddb, emr-goodi es, emr-kinesis, emr-s3-di st-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-ma pred, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server, hbase-hma ster, hbase-client, hbase-reg ion-server, hbase-rest-server, hbase-thrift-server, hbase-ope rator-tools, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-7.6.0	2.6.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-7.5.0	2.5.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-7.4.0	2.5.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-7.3.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-7.2.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server
emr-5.36.2	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-7.1.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-7.0.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-6.15.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-6.14.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-6.13.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-6.12.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-6.11.1	2.4.15	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-6.11.0	2.4.15	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-6.10.1	2.4.15	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-6.10.0	2.4.15	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-6.9.1	2.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server
emr-6.9.0	2.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-6.8.1	2.4.12	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server
emr-6.8.0	2.4.12	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-6.7.0	2.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server
emr-5.36.1	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.36.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.6.0	2.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.35.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.5.0	2.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-6.4.0	2.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.3.1	2.2.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-6.3.0	2.2.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.2.1	2.2.6-amzn-0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-6.2.0	2.2.6-amzn-0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.1.1	2.2.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-6.1.0	2.2.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.0.1	2.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-6.0.0	2.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.34.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.33.1	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.33.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.32.1	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.32.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.31.1	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.31.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.30.2	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.30.1	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.30.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.29.0	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.28.1	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.28.0	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.27.1	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.27.0	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.26.0	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.25.0	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.24.1	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.24.0	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.23.1	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.23.0	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.22.0	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.21.2	1.4.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.21.1	1.4.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.21.0	1.4.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.20.1	1.4.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.20.0	1.4.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.19.1	1.4.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.19.0	1.4.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.18.1	1.4.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.18.0	1.4.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.17.2	1.4.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.17.1	1.4.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.17.0	1.4.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.16.1	1.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.16.0	1.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.15.1	1.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.15.0	1.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.14.2	1.4.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.14.1	1.4.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.14.0	1.4.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.13.1	1.4.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.13.0	1.4.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.12.3	1.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.12.2	1.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.12.1	1.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.12.0	1.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.11.4	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.11.3	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.11.2	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.11.1	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.11.0	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.10.1	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.10.0	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.9.1	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.9.0	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.8.3	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.8.2	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.8.1	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.8.0	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.7.1	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.7.0	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.6.1	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.6.0	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.5.4	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.5.3	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.5.2	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.5.1	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.5.0	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.4.1	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.4.0	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.3.2	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.3.1	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.3.0	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.2.3	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.2.2	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.2.1	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.2.0	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.1.1	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.1.0	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.0.3	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.0.2	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.0.1	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-5.0.0	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.9.6	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-4.9.5	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.9.4	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-4.9.3	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.9.2	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-4.9.1	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.8.5	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-4.8.4	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.8.3	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-4.8.2	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.8.1	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-4.8.0	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.7.4	1.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-4.7.3	1.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.7.2	1.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-4.7.1	1.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.7.0	1.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	HBase Versão	Componentes instalados com HBase
emr-4.6.1	1.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.6.0	1.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Apache HCatalog

HCatalog é uma ferramenta que permite acessar tabelas de metastore do Hive no Pig, no Spark SQL e/ou em aplicativos personalizados. MapReduce HCatalog tem uma interface REST e um cliente de linha de comando que permitem criar tabelas ou realizar outras operações. Em seguida, você escreve seus aplicativos para acessar as tabelas usando HCatalog bibliotecas. Para obter mais informações, consulte Usando HCatalog. HCatalog é incluído nas versões 4.4.0 e posteriores do Amazon EMR.

HCatalog no Amazon EMR versão 5.8.0 e versões posteriores é compatível com o uso do Catálogo de Dados do AWS Glue como o metastore do Hive. Para obter mais informações, consulte Using AWS Glue Data Catalog as the metastore for Hive.

A tabela a seguir lista a versão HCatalog incluída na versão mais recente da série 6.x do Amazon EMR, além dos componentes com os quais o Amazon EMR instala. HCatalog

Para obter a versão dos componentes instalados HCatalog nessa versão, consulte Release 6.15.0 Component Versions.

HCatalog informações de versão do emr-6.15.0

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-6.15.0	HCatalog 3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

A tabela a seguir lista a versão HCatalog incluída na versão mais recente da série 5.x do Amazon EMR, além dos componentes com os quais o Amazon EMR instala. HCatalog

Para obter a versão dos componentes instalados HCatalog nessa versão, consulte Release 5.36.2 Component Versions.

HCatalog informações de versão do emr-5.36.2

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.36.2	HCatalog 2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Tópicos

- Criando um cluster com HCatalog
- Usando HCatalog
- Exemplo: criar uma HCatalog tabela e gravar nela usando o Pig
- HCatalog histórico de versões

Criando um cluster com HCatalog

Embora HCatalog esteja incluído no projeto Hive, você deve instalá-lo como seu próprio aplicativo.

Iniciar um cluster com HCatalog o uso do console

O procedimento a seguir cria um cluster com HCatalog instalado. Para obter mais informações sobre como criar clusters usando o console, incluindo as Opções avançadas, consulte Plan and configure clusters no Guia de gerenciamento do Amazon EMR.

1. Abra o console do Amazon EMR em https://console.aws.amazon.com /emr.
2. Escolha Create cluster (Criar cluster) para usar Quick Create.
3. Para o campo Software Configuration (Configuração do software), selecione Amazon versão emr-4.4.0 ou posterior.
4. No campo Selecionar aplicativos, escolha Todos os aplicativos ou HCatalog.
5. Selecione outras opções conforme necessário e escolha Create cluster (Criar cluster).

Para iniciar um cluster HCatalog usando o AWS CLI

- Crie o cluster usando o seguinte comando:

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr create-cluster --name "Cluster with Hcat" --release-label emr-7.9.0 \
--applications Name=HCatalog --ec2-attributes KeyName=myKey \
--instance-type m5.xlarge --instance-count 3 --use-default-roles
```

Usando HCatalog

Você pode usar HCatalog em vários aplicativos que usam o metastore Hive. Os exemplos nesta seção mostram como criar uma tabela e usá-la no contexto do Pig e do Spark SQL.

Desabilitar a gravação direta ao usar HCatalog HStorer

Sempre que uma aplicação usar o HCatStorer para gravar em uma HCatalog tabela armazenada no Amazon S3, desabilite o atributo de gravação direta do Amazon EMR. Por exemplo,

desativar a gravação direta ao usar o `STORE` comando Pig ou ao executar trabalhos do Sqoop que gravam HCatalog tabelas no Amazon S3. Você pode desativar o recurso de gravação direta definindo as configurações `mapred.output.direct.NativeS3FileSystem` e `mapred.output.direct.EmrFileSystem` como `false`. O exemplo a seguir demonstra como definir essas configurações usando o Java.

```
Configuration conf = new Configuration();
conf.set("mapred.output.direct.NativeS3FileSystem", "false");
conf.set("mapred.output.direct.EmrFileSystem", "false");
```

Criar uma tabela usando a HCat CLI e usar esses dados no Pig

Crie o seguinte script, `impressions.q`, no seu cluster:

```
CREATE EXTERNAL TABLE impressions (
    requestBeginTime string, adId string, impressionId string, referrer string,
    userAgent string, userCookie string, ip string
  )
  PARTITIONED BY (dt string)
  ROW FORMAT
    serde 'org.apache.hive.hcatalog.data.JsonSerDe'
    with serdeproperties ( 'paths'='requestBeginTime, adId, impressionId, referrer,
 userAgent, userCookie, ip' )
  LOCATION 's3://[your region].elasticmapreduce/samples/hive-ads/tables/impressions/';
ALTER TABLE impressions ADD PARTITION (dt='2009-04-13-08-05');
```

Execute o script usando a HCat CLI:

```
% hcat -f impressions.q
Logging initialized using configuration in file:/etc/hive/conf.dist/hive-
log4j.properties
OK
Time taken: 4.001 seconds
OK
Time taken: 0.519 seconds
```

Abra o shell do Grunt e acesse os dados em `impressions`:

```
% pig -useHCatalog -e "A = LOAD 'impressions' USING
 org.apache.hive.hcatalog.pig.HCatLoader();
B = LIMIT A 5;
```

```
dump B;"
<snip>
(1239610346000,m9nwdo67Nx6q2kI25qt5On7peICfUM,omkxkaRpNhGPDucAiBErSh1cs0MThC,cartoonnetwork.com
 (compatible; MSIE 7.0; Windows NT 6.0; FunWebProducts; GTB6; SLCC1; .NET CLR
 2.0.50727; Media Center PC
 5.0; .NET,wcVWWTascoPbGt6bdqDbuWTPPHgOPs,69.191.224.234,2009-04-13-08-05)
(1239611000000,NjriQjdODgWBKnkGJUP6GNTbDeK4An,AWtXPkfaWGOaNeL9OOsFU8Hcj6eLHt,cartoonnetwork.com
 (compatible; MSIE 7.0; Windows NT 5.1; GTB6; .NET CLR
 1.1.4322),OaMU1F2gE4CtADVHAbKjjRRks5kIgg,57.34.133.110,2009-04-13-08-05)
(1239610462000,Irpv3oiu0I5QNQiwSSTIshrLdo9cM1,i1LDq44LRSJF0hbmhB8Gk7k9gMWtBq,cartoonnetwork.com
 (compatible; MSIE 6.0; Windows NT 5.2; SV1; .NET CLR 1.1.4322;
 InfoPath.1),QSb3wkLR4JAIut4Uq6FNFQIR1rCVwU,42.174.193.253,2009-04-13-08-05)
(1239611007000,q2Awfnpe0JAvhInaIp0VGx9KTs0oPO,s3HvTflPB8JIE0IuM6hOEebWWpOtJV,cartoonnetwork.com
 (compatible; MSIE 6.0; Windows NT 5.2; SV1; .NET CLR 1.1.4322;
 InfoPath.1),QSb3wkLR4JAIut4Uq6FNFQIR1rCVwU,42.174.193.253,2009-04-13-08-05)
(1239610398000,c362vpAB0soPKGHRS43cj6TRwNeOGn,jeas5nXbQInGAgFB8jlkhnprN6cMw7,cartoonnetwork.com
 (compatible; MSIE 8.0; Windows NT 5.1; Trident/4.0; GTB6; .NET CLR
 1.1.4322),k96n5PnUmwHKfiUI0TFP0TNMfADgh9,51.131.29.87,2009-04-13-08-05)
7120 [main] INFO  org.apache.pig.Main  - Pig script completed in 7 seconds and 199
 milliseconds (7199 ms)
16/03/08 23:17:10 INFO pig.Main: Pig script completed in 7 seconds and 199 milliseconds
 (7199 ms)
```

Acessar a tabela usando o Spark SQL

Este exemplo cria um Spark a DataFrame partir da tabela criada no primeiro exemplo e mostra as primeiras 20 linhas:

```
% spark-shell --jars /usr/lib/hive-hcatalog/share/hcatalog/hive-hcatalog-core-1.0.0-
amzn-3.jar
<snip>
scala> val hiveContext = new org.apache.spark.sql.hive.HiveContext(sc);
scala> val df = hiveContext.sql("SELECT * FROM impressions")
scala> df.show()
<snip>
16/03/09 17:18:46 INFO DAGScheduler: ResultStage 0 (show at <console>:32) finished in
 10.702 s
16/03/09 17:18:46 INFO DAGScheduler: Job 0 finished: show at <console>:32, took
 10.839905 s
+---------------+-------------------+-------------------+-----------------
+-------------------+-------------------+-------------+---------------+
|requestbegintime|               adid|       impressionid|         referrer|
   useragent|         usercookie|            ip|             dt|
```

```
+---------------+-------------------+-------------------+-----------------
+-------------------+-------------------+-------------+---------------+
|   1239610346000|m9nwdo67Nx6q2kI25...|omkxkaRpNhGPDucAi...|cartoonnetwork.com|
Mozilla/4.0 (comp...|wcVWWTascoPbGt6bd...|69.191.224.234|2009-04-13-08-05|
|   1239611000000|NjriQjdODgWBKnkGJ...|AWtXPkfaWGOaNeL9O...|cartoonnetwork.com|
Mozilla/4.0 (comp...|OaMU1F2gE4CtADVHA...| 57.34.133.110|2009-04-13-08-05|
|   1239610462000|Irpv3oiu0I5QNQiwS...|i1LDq44LRSJF0hbmh...|cartoonnetwork.com|
Mozilla/4.0 (comp...|QSb3wkLR4JAIut4Uq...|42.174.193.253|2009-04-13-08-05|
|   1239611007000|q2Awfnpe0JAvhInaI...|s3HvTflPB8JIE0IuM...|cartoonnetwork.com|
Mozilla/4.0 (comp...|QSb3wkLR4JAIut4Uq...|42.174.193.253|2009-04-13-08-05|
|   1239610398000|c362vpAB0soPKGHRS...|jeas5nXbQInGAgFB8...|cartoonnetwork.com|
Mozilla/4.0 (comp...|k96n5PnUmwHKfiUI0...|  51.131.29.87|2009-04-13-08-05|
|   1239610600000|cjBTpruoaiEtqLuMX...|XwlohBSs8Ipxs1bRa...|cartoonnetwork.com|
Mozilla/4.0 (comp...|k96n5PnUmwHKfiUI0...|  51.131.29.87|2009-04-13-08-05|
|   1239610804000|Ms3eJHNAEItpxvimd...|4SIj4pGmgVLl625BD...|cartoonnetwork.com|
Mozilla/4.0 (comp...|k96n5PnUmwHKfiUI0...|  51.131.29.87|2009-04-13-08-05|
|   1239610872000|h5bccHX6wJReDi1jL...|EFAWIiBdVfnxwAMWP...|cartoonnetwork.com|
Mozilla/4.0 (comp...|k96n5PnUmwHKfiUI0...|  51.131.29.87|2009-04-13-08-05|
|   1239610365000|874NBpGmxNFfxEPKM...|xSvE4XtGbdtXPF2Lb...|cartoonnetwork.com|
Mozilla/5.0 (Maci...|eWDEVVUphlnRa273j...| 22.91.173.232|2009-04-13-08-05|
|   1239610348000|X8gISpUTSqh1A5reS...|TrFblGT99AgE75vuj...|       corriere.it|
Mozilla/4.0 (comp...|tX1sMpnhJUhmAF7AS...|   55.35.44.79|2009-04-13-08-05|
|   1239610743000|kbKreLWB6QVueFrDm...|kVnxx9Ie2i3OLTxFj...|       corriere.it|
Mozilla/4.0 (comp...|tX1sMpnhJUhmAF7AS...|   55.35.44.79|2009-04-13-08-05|
|   1239610812000|9lxOSRpEi3bmEeTCu...|1B2sff99AEIwSuLVV...|       corriere.it|
Mozilla/4.0 (comp...|tX1sMpnhJUhmAF7AS...|   55.35.44.79|2009-04-13-08-05|
|   1239610876000|lijjmCf2kuxfBTnjL...|AjvufgUtakUFcsIM9...|       corriere.it|
Mozilla/4.0 (comp...|tX1sMpnhJUhmAF7AS...|   55.35.44.79|2009-04-13-08-05|
|   1239610941000|t8t8trgjNRPIlmxuD...|agu2u2TCdqWP08rAA...|       corriere.it|
Mozilla/4.0 (comp...|tX1sMpnhJUhmAF7AS...|   55.35.44.79|2009-04-13-08-05|
|   1239610490000|OGRLPVNGxiGgrCmWL...|mJg2raBUpPrC8OlUm...|       corriere.it|
Mozilla/4.0 (comp...|r2k96t1CNjSU9fJKN...|   71.124.66.3|2009-04-13-08-05|
|   1239610556000|OnJID12x0RXKPUgrD...|P7Pm2mPdW6wO8KA3R...|       corriere.it|
Mozilla/4.0 (comp...|r2k96t1CNjSU9fJKN...|   71.124.66.3|2009-04-13-08-05|
|   1239610373000|WflsvKIgOqfIE5KwR...|TJHd1VBspNcua0XPn...|       corriere.it|
Mozilla/5.0 (Maci...|fj2L1ILTFGMfhdrt3...| 75.117.56.155|2009-04-13-08-05|
|   1239610768000|4MJR0XxiVCU1ueXKV...|1OhGWmbvKf8ajoU8a...|       corriere.it|
Mozilla/5.0 (Maci...|fj2L1ILTFGMfhdrt3...| 75.117.56.155|2009-04-13-08-05|
|   1239610832000|gWIrpDiN57i3sHatv...|RNL4C7xPi3tdar2Uc...|       corriere.it|
Mozilla/5.0 (Maci...|fj2L1ILTFGMfhdrt3...| 75.117.56.155|2009-04-13-08-05|
|   1239610789000|pTne9k62kJ14QViXI...|RVxJVIQousjxUVI3r...|        pixnet.net|
Mozilla/5.0 (Maci...|1bGOKiBD2xmui9OkF...| 33.176.101.80|2009-04-13-08-05|
+---------------+-------------------+-------------------+-----------------
+-------------------+-------------------+-------------+---------------+
```

```
only showing top 20 rows


scala>
```

Exemplo: criar uma HCatalog tabela e gravar nela usando o Pig

Você pode criar uma HCatalog tabela e usar o Apache Pig para gravar nela por meio do HCat Storer usando uma fonte de dados no Amazon S3. HCatalog exige que você desative a gravação direta ou a operação falhará silenciosamente. Defina as configurações `mapred.output.direct.NativeS3FileSystem` e `mapred.output.direct.EmrFileSystem` como `false` usando a classificação `mapred-site` ou manualmente no shell do Grunt. O exemplo a seguir mostra uma tabela criada usando a HCat CLI, seguida por comandos executados no shell do Grunt para preencher a tabela a partir de um arquivo de dados de exemplo no Amazon S3.

Para executar esse exemplo, conecte-se ao nó principal usando SSH.

Crie um arquivo de HCatalog script,`wikicount.q`, com o conteúdo a seguir, que cria uma HCatalog tabela chamada`wikicount`.

```
CREATE EXTERNAL TABLE IF NOT EXISTS wikicount(
col1 string,
col2 bigint
)
ROW FORMAT DELIMITED FIELDS TERMINATED BY '\001'
STORED AS ORC
LOCATION 's3://amzn-s3-demo-bucket/hcat/wikicount';
```

Use um comando HCat CLI para executar o script a partir do arquivo.

```
hcat -f wikicount.q
```

Em seguida, inicie o shell do Grunt com a opção `-useHCatalog`, defina as configurações para desativar a gravação direta, carregue os dados de um local do S3 e, em seguida, grave os resultados na tabela wikicount.

```
pig -useHCatalog
```

```
SET mapred.output.direct.NativeS3FileSystem false;
SET mapred.output.direct.EmrFileSystem false;
A = LOAD 's3://support.elasticmapreduce/training/datasets/wikistats_tiny/' USING
 PigStorage(' ') AS (Site:chararray, page:chararray, views:int, total_bytes:long);
B = GROUP A BY Site;
C = FOREACH B GENERATE group as col1, COUNT(A) as col2;
STORE C INTO 'wikicount' USING org.apache.hive.hcatalog.pig.HCatStorer();
```

HCatalog histórico de versões

A tabela a seguir lista a versão HCatalog incluída em cada versão do Amazon EMR, junto com os componentes instalados com a aplicação. Para obter as versões dos componentes em cada versão, consulte a seção Versão de componentes da sua versão em Versões 7.x do Amazon EMR, Versões de lançamento 6.x do Amazon EMR ou Versões de lançamento 5.x do Amazon EMR.

HCatalog informações sobre a versão

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-7.9.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-7.8.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode,

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
		hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-7.7.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-7.6.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-7.5.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-7.4.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-7.3.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-7.2.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.36.2	2,3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-7.1.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-7.0.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-6.15.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.14.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-6.13.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.12.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-6.11.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.11.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-6.10.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.10.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-6.9.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.9.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-6.8.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.8.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-6.7.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.36.1	2,3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.36.0	2,3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.6.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.35.0	2,3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.5.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-6.4.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.3.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-6.3.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.2.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-6.2.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.1.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-6.1.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.0.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-6.0.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.34.0	2.3.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.33.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.33.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.32.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.32.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.31.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.31.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.30.2	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.30.1	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.30.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.29.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.28.1	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.28.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.27.1	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.27.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.26.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.25.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.24.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.24.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.23.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.23.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.22.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.21.2	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.21.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.21.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.20.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.20.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.19.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.19.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.18.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.18.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.17.2	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.17.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.17.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.16.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.16.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.15.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.15.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.14.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.14.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.14.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.13.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.13.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.12.3	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.12.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.12.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.12.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.11.4	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.11.3	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.11.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.11.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.11.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.10.1	2.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.10.0	2.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.9.1	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.9.0	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.8.3	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.8.2	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.8.1	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.8.0	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.7.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.7.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.6.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.6.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.5.4	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.5.3	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.5.2	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.5.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.5.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.4.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.4.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.3.2	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.3.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.3.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.2.3	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.2.2	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.2.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.2.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.1.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.1.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.0.3	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.0.2	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.0.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-5.0.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.9.6	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-4.9.5	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.9.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-4.9.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.9.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-4.9.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.8.5	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-4.8.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.8.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-4.8.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.8.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-4.8.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.7.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-4.7.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.7.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-4.7.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.7.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-4.6.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, hive-metastore-server, mysql-server
emr-4.6.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, hive-metastore-server, mysql-server

Rótulo de versão do Amazon EMR	HCatalog Versão	Componentes instalados com HCatalog
emr-4.5.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, hive-metastore-server, mysql-server
emr-4.4.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, hive-metastore-server, mysql-server

Apache Hive

Hive é um data warehouse e um pacote de análises de código aberto executado além de um cluster do Hadoop. Os scripts do Hive usam uma linguagem semelhante a SQL chamada Hive QL (query language, linguagem de consulta) que abstrai modelos de programação e dá suporte a interações de data warehouse típicas. O Hive permite que você evite as complexidades de escrever trabalhos Tez com base em gráficos acíclicos direcionados (DAGs) ou MapReduce programas em uma linguagem de computador de nível inferior, como Java.

O Hive amplia o paradigma do SQL incluindo formatos de serialização. Você também pode personalizar o processamento de consultas com a criação de um esquema de tabela que corresponda a seus dados, sem tocar nos próprios dados. Enquanto o SQL oferece suporte a tipos de valor primitivos, como datas, números e strings, os valores de tabelas do Hive são elementos estruturados, como objetos JSON, qualquer tipo de dados definido pelo usuário ou qualquer função escrita em Java.

Para obter mais informações sobre o Hive, consulte http://hive.apache.org/.

A tabela a seguir lista a versão do Hive incluída na versão mais recente da série 7.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Hive.

Para a versão dos componentes instalados com o Hive nesta versão, consulte Versões de componentes da versão 7.9.0.

Informações sobre a versão do Hive para o emr-7.9.0

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-7.9.0	Hive 3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-ya

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
		rn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

A tabela a seguir lista a versão do Hive incluída na versão mais recente da série 6.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Hive.

Para obter a versão dos componentes instalados com o Hive nessa versão, consulte Release 6.15.0 Component Versions.

Informações de versão do Hive para o emr-6.15.0

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-6.15.0	Hive 3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

A tabela a seguir lista a versão do Hive incluída na versão mais recente da série 5.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Hive.

Para obter a versão dos componentes instalados com o Hive nesse lançamento, consulte Release 5.36.2 Component Versions.

Informações de versão do Hive para o emr-5.36.2

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.36.2	Hive 2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Desde a versão 5.18.0 do Amazon EMR, você pode usar o repositório de artefatos do Amazon EMR para criar o código de trabalho em comparação com as versões exatas de bibliotecas e dependências disponíveis com versões específicas do Amazon EMR. Para obter mais informações, consulte Verificar dependências usando o repositório de artefatos do Amazon EMR.

Tópicos

- Diferenças e considerações sobre o Hive no Amazon EMR
- Configurar um metastore externo para o Hive
- Usar o driver JDBC do Hive
- Melhorar a performance do Hive
- Usar o Live Long and Process (LLAP) do Hive

- Criptografia no Hive
- Histórico de versões do Hive

Diferenças e considerações sobre o Hive no Amazon EMR

Diferenças entre o Apache Hive no Amazon EMR e o Apache Hive

Esta seção descreve as diferenças entre o Hive no Amazon EMR e as versões padrão do Hive disponíveis em http://svn.apache. org/viewvc/hive/branches/.

Autorização do Hive

O Amazon EMR é compatível com a autorização do Hive para o HDFS, mas não para o EMRFS e o Amazon S3. Os clusters do Amazon EMR são executados com a autorização desabilitada por padrão.

Comportamento de mesclagem de arquivos do Hive com o Amazon S3

O Apache Hive mesclará arquivos pequenos no final de um trabalho somente mapa se `hive.merge.mapfiles` for "true", e a mesclagem apenas será acionada se o tamanho médio de saída do trabalho for menor do que a configuração `hive.merge.smallfiles.avgsize`. O Hive no Amazon EMR tem exatamente o mesmo comportamento se o caminho de saída final está no HDFS. Se o caminho de saída estiver no Amazon S3, o parâmetro `hive.merge.smallfiles.avgsize` será ignorado. Nessa situação, a tarefa de mesclagem sempre será acionada se `hive.merge.mapfiles` estiver definido como `true`.

Transações ACID e o Amazon S3

O Amazon EMR 6.1.0 e versões posteriores oferecem suporte às transações Hive ACID (atomicidade, consistência, isolamento, durabilidade) para que esteja em conformidade com as propriedades ACID de um banco de dados. Com esse atributo, você pode executar as operações INSERT, UPDATE, DELETE e MERGE em tabelas gerenciadas do Hive com dados no Amazon Simple Storage Service (Amazon S3).

LLAP (Live Long and Process) do Hive

A funcionalidade LLAP adicionada na versão 2.0 do Apache Hive padrão não é compatível com o Hive 2.1.0 na versão 5.0 do Amazon EMR.

As versões 6.0.0 e posteriores do Amazon EMR são compatíveis com a funcionalidade Live Long and Process (LLAP) do Hive. Para obter mais informações, consulte Usar o LLAP do Hive.

Diferenças no Hive entre as versões 4.x e 5.x do Amazon EMR

Esta seção discute as diferenças que devem ser consideradas antes da migração de uma implementação do Hive versão 1.0.0 na versão 4.x do Amazon EMR para o Hive 2.x na versão 5.x do Amazon EMR.

Considerações e diferenças operacionais

- Suporte adicionado para transações ACID (atomicidade, consistência, isolamento e durabilidade): essa diferença entre o Hive 1.0.0 no Amazon EMR 4.x e o Apache Hive padrão foi eliminada.
- Gravações diretas no Amazon S3 eliminadas: essa diferença entre o Hive 1.0.0 no Amazon EMR e o Apache Hive padrão foi eliminada. O Hive 2.1.0 na versão 5.x do Amazon EMR agora cria, lê e grava arquivos temporários armazenados no Amazon S3. Como resultado, para ler e gravar na mesma tabela, você não precisa mais criar uma tabela temporária no sistema de arquivos HDFS local do cluster como solução alternativa. Se você usar buckets com versionamento, certifique-se de gerenciar esses arquivos temporários conforme descrito abaixo.
- Gerenciar arquivos temporários ao serem usados buckets do Amazon S3 com versionamento: quando você executa consultas do Hive em que o destino dos dados gerados é o Amazon S3, muitos arquivos e diretórios e temporários são criados. Este é o novo comportamento, conforme descrito anteriormente. Se você usar buckets do S3 com versionamento, esses arquivos temporários sobrecarregarão o Amazon S3 e resultarão em custos caso não sejam excluídos. Ajuste suas regras de ciclo de vida para que os dados com um prefixo `/_tmp` sejam excluídos após um curto período, como cinco dias. Consulte Especificar uma configuração de ciclo de vida para obter mais informações.
- Log4j atualizado para log4j 2: se você usa log4j, talvez precise alterar sua configuração de registro em log devido a este upgrade. Para obter detalhes, consulte Apache log4j 2.

Diferenças e considerações de desempenho

- Diferenças de performance com o Tez: com a versão 5.x do Amazon EMR, o Tez é o mecanismo de execução padrão para o Hive em vez do. MapReduce O Tez fornece melhor desempenho para a maioria dos fluxos de trabalho.
- Tabelas com muitas partições: consultas que geram um grande número de partições dinâmicas podem falhar e consultas que selecionam tabelas com muitas partições podem demorar mais do

que o esperado para serem executadas. Por exemplo, um seleção de 100.000 partições pode demorar 10 minutos ou mais.

Atributos adicionais do Hive no Amazon EMR

O Amazon EMR amplia o Hive com novos atributos compatíveis com a integração do Hive com outros AWS serviços da, como a capacidade de ler e gravar no Amazon Simple Storage Service (Amazon S3) e no DynamoDB.

Variáveis no Hive

Você pode incluir variáveis nos seus scripts usando o sinal de cifrão e chaves.

```
add jar ${LIB}/jsonserde.jar
```

Você transmite os valores dessas variáveis ao Hive na linha de comando usando o parâmetro `-d`, como no exemplo a seguir:

```
-d LIB=s3://elasticmapreduce/samples/hive-ads/lib
```

Você também pode transmitir os valores em etapas que executam scripts Hive.

Para transmitir valores variáveis em etapas do Hive usando o console

1. Abra o console do Amazon EMR em https://console.aws.amazon.com /emr.
2. Selecione Criar cluster.
3. Na seção Steps (Etapas), para Add Step (Adicionar etapa), escolha Hive Program (Programa Hive) na lista e depois Configure and add (Configurar e adicionar).
4. Na caixa de diálogo Add Step (Adicionar etapa), especifique os parâmetros usando a tabela a seguir como guia e escolha Add (Adicionar).

Campo	Ação
Localização do script no S3*	Especificar o URI no qual o script reside no Amazon S3. O valor deve estar no formato*BucketName* /*path*/*ScriptName* . Por exemplo: `s3://elasticmapreduce/samples/hive-ads/libs/response-time-stats.q` .
Localização do S3 de entrada	Opcionalmente, especifique o URI no qual os arquivos de entrada residem no Amazon S3. O valor deve estar no formato*BucketName* /*path*/. Se especificadas, essas informações serão transmitidas ao script do Hive como um parâmetro chamado `INPUT`. Por exemplo: `s3://elasticmapreduce/samples/hive-ads/tables/` .
Local do S3 de saída	Opcionalmente, especifique o URI no qual você deseja a saída armazenada no Amazon S3. O valor deve estar no formato*BucketName* /*path*. Se especificadas, essas informações serão transmitidas ao script do Hive como um parâmetro chamado `0UTPUT`. Por exemplo: `s3://amzn-s3-demo-bucket/hive-ads/output/` .
Argumentos	Opcionalmente, insira uma lista de argumentos (strings separadas por espaços) para transmitir ao Hive. Se você tiver definido uma variável de caminho no script Hive chamada ${SAMPLE}, por exemplo: `CREATE EXTERNAL TABLE logs (requestBeginTime STRING, requestEndTime STRING, hostname STRING) PARTITIONED BY (dt STRING) \ ROW FORMAT serde 'com.amazon.elasticmapreduce.JsonSerde' WITH SERDEPROPERTIES ( 'paths'='requestBeginTime, requestEndTime, hostname' ) LOCATION '${SAMPLE}/tables/impressions';` Para transmitir um valor para a variável, digite o seguinte na janela Arguments (Argumentos): `-d SAMPLE=s3://elasticmapreduce/samples/hive-ads/` .

Campo	Ação
Ação em caso de falha	Isso determina o que faz o cluster em resposta a qualquer erro. Os valores possíveis para essa configuração são: • Terminate cluster (Encerrar o cluster): se a etapa falhar, encerrar o cluster. Se o cluster tiver proteção contra encerramento habilitado E o keep alive habilitado, ele não será encerrado. • Cancel and wait (Cancelar e aguardar): se a etapa falhar, cancele as etapas restantes. Se o cluster tiver o keep alive habilitado, ele não será encerrado. • Continue (Continuar): se a etapa falhar, continue na próxima etapa.

5. Selecione os valores conforme necessário e escolha Create cluster (Criar cluster).

Para passar valores variáveis para as etapas do Hive usando o AWS CLI

Para passar valores variáveis para as etapas do Hive usando o AWS CLI, use o `--steps` parâmetro e inclua uma lista de argumentos.

-
> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.9.0 \
--applications Name=Hive Name=Pig --use-default-roles --ec2-attributes
 KeyName=myKey --instance-type m5.xlarge --instance-count 3 \
--steps Type=Hive,Name="Hive Program",ActionOnFailure=CONTINUE,Args=[-f,s3://
elasticmapreduce/samples/hive-ads/libs/response-time-stats.q,-d,INPUT=s3://
elasticmapreduce/samples/hive-ads/tables,-d,OUTPUT=s3://amzn-s3-demo-bucket/hive-
ads/output/,-d,SAMPLE=s3://elasticmapreduce/samples/hive-ads/]
```

Para obter mais informações sobre o uso dos comandos do Amazon EMR no AWS CLI, consulte. https://docs.aws.amazon.com/cli/latest/reference/emr

Para transmitir valores variáveis em etapas do Hive usando o SDK Java

- O exemplo a seguir demonstra como transmitir variáveis em etapas usando o SDK. Para obter mais informações, consulte Classe StepFactory na Referência AWS SDK para Java da API.

```
StepFactory stepFactory = new StepFactory();

   StepConfig runHive = new StepConfig()
     .withName("Run Hive Script")
     .withActionOnFailure("TERMINATE_JOB_FLOW")
     .withHadoopJarStep(stepFactory.newRunHiveScriptStep(“s3://amzn-s3-demo-bucket/
script.q”,
      Lists.newArrayList(“-d”,”LIB= s3://elasticmapreduce/samples/hive-ads/lib”));
```

Consultas do Hive no Amazon EMR para acomodar esquemas parciais do DynamoDB

O Hive no Amazon EMR oferece a máxima flexibilidade na consulta a tabelas do DynamoDB ao permitir que você especifique um subconjunto de colunas no qual pode filtrar dados, em vez de exigir que sua consulta inclua todas as colunas. Essa técnica de consulta de esquema parcial é eficaz quando você tem uma esquema de banco de dados esparso e deseja filtrar registros com base em algumas colunas, por exemplo, em carimbos de data/hora.

O exemplo a seguir mostra como usar uma consulta do Hive para:

- Crie uma tabela do DynamoDB.
- Selecione um subconjunto de itens (linhas) no DynamoDB e restrinja ainda mais os dados para determinadas colunas.
- Copie os dados resultantes para o Amazon S3.

```
DROP TABLE dynamodb;
DROP TABLE s3;

CREATE EXTERNAL TABLE dynamodb(hashKey STRING, recordTimeStamp BIGINT, fullColumn
 map<String, String>)
    STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
    TBLPROPERTIES (
     "dynamodb.table.name" = "myTable",
     "dynamodb.throughput.read.percent" = ".1000",
     "dynamodb.column.mapping" = "hashKey:HashKey,recordTimeStamp:RangeKey");
```

```
CREATE EXTERNAL TABLE s3(map<String, String>)
     ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
     LOCATION 's3://bucketname/path/subpath/';

INSERT OVERWRITE TABLE s3 SELECT item fullColumn FROM dynamodb WHERE recordTimeStamp <
 "2012-01-01";
```

A tabela a seguir mostra a sintaxe de consulta para a seleção de qualquer combinação de itens do DynamoDB.

Exemplo de consulta	Descrição do resultado
SELECIONE * DE*table_name* ;	Seleciona todos os itens (linhas) de uma determinada tabela e inclui dados de todas as colunas disponíveis para esses itens.
SELECIONE * DE *table_name* ONDE *field_name* =*value*;	Seleciona alguns itens (linhas) de uma determina da tabela e inclui dados de todas as colunas disponíveis para esses itens.
SELECIONE*column1_name* ,*column2_name* , *column3_name* DE*table_name* ;	Seleciona todos os itens (linhas) de uma determinada tabela e inclui dados de algumas colunas disponíveis para esses itens.
SELECIONE*column1_name* ,*column2_name* , *column3_name* DE *table_name* ONDE *field_name* =*value*;	Seleciona alguns itens (linhas) de uma determina da tabela e inclui dados de algumas colunas disponíveis para esses itens.

Copiar dados entre tabelas do DynamoDB em diferentes regiões da AWS

O Hive no Amazon EMR fornece uma propriedade `dynamodb.region` que você pode definir para cada tabela do DynamoDB. Quando `dynamodb.region` é definido de forma diferente em duas tabelas, quaisquer dados copiados entre as tabelas ocorrem automaticamente entre as regiões especificadas.

O exemplo a seguir mostra como criar uma tabela do DynamoDB com um script do Hive que define a propriedade `dynamodb.region`:

> **Note**
>
> As propriedades da região por tabela substituem as propriedades globais do Hive.

```
CREATE EXTERNAL TABLE dynamodb(hashKey STRING, recordTimeStamp BIGINT, map<String,
 String> fullColumn)
    STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
    TBLPROPERTIES (
     "dynamodb.table.name" = "myTable",
     "dynamodb.region" = "eu-west-1",
     "dynamodb.throughput.read.percent" = ".1000",
     "dynamodb.column.mapping" = "hashKey:HashKey,recordTimeStamp:RangeKey");
```

Definir valores de throughput do DynamoDB por tabela

O Hive no Amazon EMR permite que você defina o DynamoDB readThroughputPercent e writeThroughputPercent as configurações para cada tabela na definição de tabelas. O script do Hive no Amazon EMR a seguir mostra como definir os valores de throughput. Para obter mais informações sobre valores de throughput do DynamoDB, consulte Especificar requisitos de leitura e gravação para tabelas.

```
CREATE EXTERNAL TABLE dynamodb(hashKey STRING, recordTimeStamp BIGINT, map<String,
 String> fullColumn)
    STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
    TBLPROPERTIES (
     "dynamodb.table.name" = "myTable",
     "dynamodb.throughput.read.percent" = ".4",
     "dynamodb.throughput.write.percent" = "1.0",
     "dynamodb.column.mapping" = "hashKey:HashKey,recordTimeStamp:RangeKey");
```

Configurar um metastore externo para o Hive

Por padrão, o Hive registra as informações do metastore em um banco de dados MySQL no sistema de arquivos do nó primário. O metastore contém uma descrição da tabela e dos dados subjacentes nos quais foi criado, incluindo os nomes de partição, os tipos de dados etc. Quando um cluster é encerrado, todos os nós de cluster são desligados, inclusive o nó primário. Quando isso acontece, os dados locais são perdidos porque os sistemas de arquivos dos nós usam armazenamento

temporário. Se precisar que o metastore seja persistido, você deverá criar um metastore externo que exista fora do cluster.

Você tem duas opções para um metastore externo:

- AWS Catálogo de Dados do Glue (somente versão 5.8.0 ou posterior do Amazon EMR).

  Para obter mais informações, consulte Usar o Catálogo de Dados do AWS Glue como metastore do Hive.
- Amazon RDS ou Amazon Aurora.

  Para obter mais informações, consulte Usar um banco de dados externo MySQL ou Amazon Aurora.

> **Note**
>
> Se você estiver usando o Hive 3 e encontrar muitas conexões com o metastore do Hive, configure o parâmetro `datanucleus.connectionPool.maxPoolSize` para ter um valor menor ou aumente o número de conexões que o servidor de banco de dados pode processar. O aumento do número de conexões se deve à forma como o Hive calcula o número máximo de conexões JDBC. Para calcular o valor ideal de performance, consulte Propriedades de configuração do Hive.

Usar o Catálogo de Dados do AWS Glue como metastore do Hive

Usando a versão 5.8.0 ou posterior do Amazon EMR, você pode configurar o Hive para usar o Catálogo de Dados do AWS Glue como metastore dele. Recomendamos essa configuração quando você precisa de um metastore persistente ou de um metastore compartilhado por diferentes clusters, serviços e aplicativos ou contas da AWS .

AWS O Glue é um serviço de extração, transformação e carregamento (ETL) totalmente gerenciado que torna fácil e econômico categorizar os dados, limpá-los, aprimorá-los e movê-los de modo confiável entre vários armazenamentos de dados. O Catálogo de Dados do AWS Glue fornece um repositório de metadados unificado em uma grande variedade de fontes e formatos de dados, integrando-se não apenas ao Amazon EMR, mas ao Amazon RDS, ao Amazon RDS, ao Amazon Redshift, ao Redshift Spectrum, ao Athena e a qualquer aplicação compatível com o metastore do Apache Hive. Os crawlers do AWS Glue podem inferir automaticamente o esquema com base nos

dados de origem no Amazon S3 e armazenar os metadados associados no Catálogo de Dados. Para obter mais informações sobre o Catálogo de Dados, consulte Preencher o Catálogo de Dados do AWS Glue no Guia do desenvolvedor do AWS Glue.

São feitas cobranças separadas pelo uso do AWS Glue. Existe uma taxa mensal para armazenar e acessar os metadados no Catálogo de Dados, uma taxa horária cobrada por minuto para trabalhos de extração, transformação e carregamento (ETL) do AWS Glue e runtime do crawler e uma taxa horária cobrada por minuto para cada endpoint de desenvolvimento provisionado. O Catálogo de Dados permite o armazenamento de até um milhão de objetos gratuitamente. Se você armazenar mais de um milhão de objetos, será cobrado uma taxa de USD$1 por cada 100.000 objetos em mais de um milhão. Um objeto no Catálogo de Dados é uma tabela, uma partição ou um banco de dados. Para obter mais informações, consulte Definição de preço do Glue.

> **Important**
>
> Se você tiver criado tabelas usando o Amazon Athena ou o Amazon Redshift Spectrum antes de 14 de agosto de 2017, os bancos de dados e as tabelas serão armazenados em um catálogo gerenciado pelo Athena, que é separado do Catálogo de Dados do Glue. AWS Para integrar o Amazon EMR a essas tabelas, faça a atualização para o Catálogo de Dados do AWS Glue. Para obter mais informações, consulte Atualização para o catálogo de dados AWS Glue no Guia do usuário do Amazon Athena.

Especificar o Catálogo de Dados do AWS Glue como metastore

Você pode especificar o Catálogo de Dados do AWS Glue como metastore usando o AWS Management Console AWS CLI, a ou a API do Amazon EMR. Ao usar a CLI ou a API, você usa a classificação de configuração do Hive para especificar o Catálogo de Dados. Além disso, com as versões 5.16.0 e posteriores do Amazon EMR, você pode usar a classificação de configuração para especificar um Catálogo de Dados em outra. Conta da AWS Ao usar o console, você pode especificar o Catálogo de Dados usando Opções avançadas ou Opções rápidas.

Console

Para especificar o AWS Glue Data Catalog como metastore do Hive com o console

1. Faça login no e abra AWS Management Console o console do Amazon EMR em https://console.aws.amazon.com /emr.

2. Em EMR, EC2 no painel de navegação esquerdo, escolha Clusters e, em seguida, escolha Criar cluster.
3. Em Pacote de aplicativos, escolha Core Hadoop ou Custom HBase. Se você personalizar o cluster, selecione Hive ou HCatalog como uma das suas aplicações.
4. Em Configurações do Catálogo de Dados do AWS Glue, marque a caixa de seleção Usar para metadados de tabelas do Hive.
5. Escolha qualquer outra opção que se aplique ao cluster.
6. Para iniciar o cluster, escolha Criar cluster.

CLI

Para especificar o Catálogo de Dados do AWS Glue como metastore do Hive usando a AWS CLI

Para obter mais informações sobre como especificar uma classificação de configuração usando a AWS CLI e a API do EMR, consulte. Configurar aplicações

- Especifique o valor para `hive.metastore.client.factory.class` usando a classificação de configuração `hive-site`, conforme mostrado neste exemplo:

```
[
  {
    "Classification": "hive-site",
    "Properties": {
      "hive.metastore.client.factory.class":
 "com.amazonaws.glue.catalog.metastore.AWSGlueDataCatalogHiveClientFactory"
    }
  }
]
```

Na versão 5.28.0, 5.28.1, 5.29.0 ou 6.x do EMR, se você estiver criando um cluster usando o Catálogo de Dados do AWS Glue como metastore, defina a. `hive.metastore.schema.verification` `false` Isso impede que o Hive and valide o esquema HCatalog do metastore em relação ao MySQL. Sem essa configuração, o grupo de instâncias principais será suspenso após a reconfiguração no Hive ou. HCatalog

```
[
  {
```

```
    "Classification": "hive-site",
    "Properties": {
      "hive.metastore.client.factory.class":
 "com.amazonaws.glue.catalog.metastore.AWSGlueDataCatalogHiveClientFactory",
      "hive.metastore.schema.verification": "false"
    }
  }
]
```

Se você já tiver um cluster na versão 5.28.0, 5.28.1 ou 5.29.0 do EMR, poderá definir o grupo de instâncias principais `hive.metastore.schema.verification` como `false` com as seguintes informações:

```
    Classification = hive-site
    Property       = hive.metastore.schema.verification
    Value          = false
```

Para especificar um Catálogo de Dados em outra AWS conta da, adicione a `hive.metastore.glue.catalogid` propriedade conforme mostrado no exemplo a seguir. Substitua *`acct-id`* pela conta da AWS do Catálogo de Dados.

```
[
  {
    "Classification": "hive-site",
    "Properties": {
      "hive.metastore.client.factory.class":
 "com.amazonaws.glue.catalog.metastore.AWSGlueDataCatalogHiveClientFactory",
      "hive.metastore.schema.verification": "false",
      "hive.metastore.glue.catalogid": "acct-id"
    }
  }
]
```

permissões do IAM

O perfil de EC2 instância de um cluster deve ter permissões do IAM para ações do AWS Glue. Além disso, se você habilitar a criptografia para objetos do Catálogo de Dados do AWS Glue, o perfil também deverá ter permissão para criptografar, descriptografar e gerar a AWS KMS key usada para criptografia.

Permissões para todas as ações do AWS Glue

Se você usa o perfil de EC2 instância padrão para o Amazon EMR, nenhuma ação será necessária. A política `AmazonElasticMapReduceforEC2Role` gerenciada anexada ao `EMR_EC2_DefaultRole` permite todas as ações necessárias do AWS Glue. No entanto, se você especificar um perfil e permissões de EC2 instância personalizados, deverá configurar as ações apropriadas do AWS Glue. Use a política gerenciada `AmazonElasticMapReduceforEC2Role` como ponto de partida. Para obter mais informações, consulte Função de serviço para EC2 instâncias de cluster (perfil de EC2 instância) no Guia de gerenciamento do Amazon EMR.

Permissões para criptografar e descriptografar o Catálogo de Dados do AWS Glue

O perfil de instância precisa de permissão para criptografar e descriptografar dados usando a chave. Você não precisa configurar essas permissões se as duas instruções a seguir se aplicarem:

- Você habilita a criptografia para objetos do Catálogo de Dados do AWS Glue usando chaves gerenciadas para o AWS Glue.
- Você usa um cluster que está na mesma que o Catálogo Conta da AWS de Dados do AWS Glue.

Caso contrário, você deverá adicionar a seguinte instrução à política de permissões anexada ao perfil de EC2 instância.

```
[
    {
        "Version": "2012-10-17",
        "Statement": [
            {
                "Effect": "Allow",
                "Action": [
                    "kms:Decrypt",
                    "kms:Encrypt",
                    "kms:GenerateDataKey"
                ],
```

```
                "Resource": "arn:aws:kms:region:acct-
id:key/12345678-1234-1234-1234-123456789012"
            }
        ]
    }
]
```

Para obter mais informações sobre criptografia do Catálogo de Dados do AWS Glue, consulte Criptografar seu catálogo de dados no Guia do desenvolvedor do AWS Glue.

Permissões baseadas em recursos

Se você usa o AWS Glue com o Hive, o Spark ou o Presto no Amazon EMR AWS , o Glue oferece suporte a políticas baseadas em recursos para controlar o acesso aos recursos do Catálogo de Dados. Esses recursos incluem bancos de dados, tabelas, conexões e funções definidas pelo usuário. Para obter mais informações, consulte Políticas baseadas em recursos no AWS Glue no Guia do desenvolvedor do AWS Glue.

Ao usar políticas baseadas em recursos para limitar o acesso ao AWS Glue de dentro do Amazon EMR, a entidade principal que você especifica na política de permissões deve ser o ARN do perfil associado ao perfil de EC2 instância que é especificado quando um cluster é criado. Por exemplo, para uma política baseada em recursos anexada a um catálogo, você pode especificar o ARN do perfil para o perfil de serviço padrão de EC2 instâncias do cluster, *EMR_EC2_DefaultRole* como a`Principal`, usando o formato mostrado no seguinte exemplo:

```
arn:aws:iam::acct-id:role/EMR_EC2_DefaultRole
```

*acct-id*Pode ser diferente do ID da conta do AWS Glue. Isso possibilita o acesso de clusters do EMR em outras contas. Você pode especificar várias entidades principais, cada uma de uma conta diferente.

Considerações ao usar o Catálogo de Dados do AWS Glue

Considere os seguintes itens ao usar o Catálogo de Dados do AWS Glue como metastore com o Hive:

- A adição de auxiliares JARs usando o shell Hive não é suportada. Como solução alternativa, use a classificação de `hive-site` configuração para definir a `hive.aux.jars.path` propriedade, que adiciona auxiliar JARs ao classpath do Hive.
- Não há suporte para transações do Hive.

- Não há suporte para renomear tabelas de dentro do AWS Glue.
- Quando você cria uma tabela do Hive sem especificar um `LOCATION`, os dados da tabela são armazenados no local especificado pela propriedade `hive.metastore.warehouse.dir`. Por padrão, esse é um local no HDFS. Se outro cluster precisar acessar a tabela, ele apresentará falha, a menos que tenha permissões adequadas para o cluster que criou a tabela. Além disso, como o armazenamento do HDFS é transitório, se o cluster for encerrado, os dados da tabela serão perdidos e a tabela precisará ser recriada. Recomendamos que você especifique a `LOCATION` no Amazon S3 ao criar uma tabela do Hive usando o Glue. AWS Como alternativa, você pode usar a classificação de configuração `hive-site` para especificar um local no Amazon S3 para `hive.metastore.warehouse.dir`, que se aplica a todas as tabelas do Hive. Se uma tabela for criada em um local do HDFS e o cluster que a criou ainda estiver em execução, você poderá atualizar o local da tabela para o Amazon S3 de dentro AWS do Glue. Para obter mais informações, consulte Como trabalhar com tabelas no console do AWS Glue no Guia do desenvolvedor do AWS Glue.
- Valores de partição contendo aspas e apóstrofos não são compatíveis, por exemplo, `PARTITION (owner="Doe's").`
- As estatísticas de coluna são têm suporte para o emr-5.31.0 e versões posteriores.
- O uso da autorização do Hive não é compatível. Como alternativa, considere usar as Políticas baseadas em recursos do AWS Glue. Para obter mais informações, consulte Usar políticas baseadas em recursos para acesso do Amazon EMR ao Catálogo de Dados do AWS Glue.
- Não há suporte para restrições do Hive.
- A Otimização baseada em custos no Hive não é compatível.
- A configuração de `hive.metastore.partition.inherit.table.properties` não é compatível.
- O uso das seguintes constantes de metastore não é compatível: `BUCKET_COUNT, BUCKET_FIELD_NAME, DDL_TIME, FIELD_TO_DIMENSION, FILE_INPUT_FORMAT, FILE_OUTPUT_FORMAT, HIVE_FILTER_FIELD_LAST_ACCESS, HIVE_FILTER_FIELD_OWNER, HIVE_FILTER_FIELD_PARAMS, IS_ARCHIVED, META_TABLE_COLUMNS, META_TABLE_COLUMN_TYPES, META_TABLE_DB, META_TABLE_LOCATION, META_TABLE_NAME, META_TABLE_PARTITION_COLUMNS, META_TABLE_SERDE, META_TABLE_STORAGE, ORIGINAL_LOCATION`.
- Quando você usa uma expressão do predicado, os valores explícitos devem estar no lado direito do operador de comparação ou as consultas podem apresentar falha.
  - Correto: `SELECT * FROM mytable WHERE time > 11`

  - Incorreto: `SELECT * FROM mytable WHERE 11 > time`
- As versões 5.32.0 e 6.3.0 e posteriores do Amazon EMR são compatíveis com o uso de funções definidas pelo usuário () em expressões do predicado. UDFs Quando são usadas versões anteriores, suas consultas podem apresentar falha devido à forma como o Hive tenta otimizar a execução da consulta.
- Tabelas temporárias não são compatíveis.
- Recomendamos a criação de tabelas usando aplicações por meio do Amazon EMR em vez de criá-las diretamente usando o AWS Glue. Criar uma tabela por meio do AWS Glue pode fazer com que campos obrigatórios estejam ausentes, causando exceções de consulta.
- Nas versões 5.20.0 ou posteriores do EMR, a remoção paralela de partições é habilitada automaticamente para o Spark e o Hive quando o Catálogo de Dados do AWS Glue é usado como metastore. Essa alteração reduz significativamente o tempo de planejamento de consultas ao executar várias solicitações em paralelo para recuperar partições. O número total de segmentos que podem ser executados simultaneamente varia entre 1 e 10. O valor padrão é 5, que é uma configuração recomendada. Você pode alterá-lo especificando a propriedade `aws.glue.partition.num.segments` na classificação de configuração `hive-site`. Se ocorrer controle de utilização, você poderá desativar o atributo alterando o valor para 1. Para obter mais informações, consulte a Estrutura de segmentos do AWS Glue.

Usar um banco de dados externo MySQL ou Amazon Aurora

Para usar um banco de dados externo MySQL ou Amazon Aurora como seu metastore do Hive, substitua os valores padrão da configuração do metastore no Hive para especificar o local do banco de dados externo, que pode ser uma instância do MySQL do Amazon RDS ou uma instância do PostgreSQL do Amazon Aurora. SQLinstance

> **Note**
>
> O Hive não oferece suporte nem impede o acesso simultâneo para gravação em tabelas da metastore. Se você compartilhar informações do metastore entre dois clusters, deverá garantir não gravar na mesma tabela do metastore simultaneamente, a menos que esteja gravando em partições diferentes da mesma tabela do metastore.

O procedimento a seguir mostra como substituir os valores de configuração padrão para a localização da metastore do Hive e iniciar um cluster usando a localização da metastore reconfigurada.

Para criar uma metastore localizada fora do cluster do EMR

1. Crie um banco de dados MySQL ou Aurora PostgreSQL. Se você usa o PostgreSQL, deve configurá-lo depois de provisionar o cluster. Somente o MySQL é compatível na criação do cluster. Para obter informações sobre as diferenças entre o Aurora MySQL e o Aurora PostgreSQL, consulte Visão geral do Amazon Aurora MySQL e Trabalhar com o Amazon Aurora PostgreSQL. Para obter informações sobre como criar um banco de dados do Amazon RDS em geral, consulte https://aws.amazon.com/rds/.
2. Modifique seus grupos de segurança para permitir conexões JDBC entre seu banco de dados e o grupo de segurança ElasticMapReduce-Master. Para obter informações sobre como modificar os grupos de segurança para obter acesso, consulte Working with Amazon EMR-managed security groups.
3. Defina valores de configuração do JDBC em `hive-site.xml`:

   > **Important**
   >
   > Se você fornecer informações confidenciais, como senhas, à API de configuração do Amazon EMR, estas informações serão exibidas para as contas que tiverem permissões suficientes. Se você teme que essas informações possam ser exibidas para outros usuários, crie o cluster com uma conta administrativa e limite os outros usuários (usuários do IAM ou aqueles com credenciais delegadas) para o acesso a serviços no cluster, criando uma função que negue permissões explicitamente para a chave de API `elasticmapreduce:DescribeCluster`.

   a. Crie um arquivo de configuração chamado `hiveConfiguration.json`, que contém edições em `hive-site.xml`, conforme mostrado no exemplo a seguir.

      *`hostname`*Substitua pelo endereço DNS da instância do Amazon RDS que executa o banco de dados *`username`* e *`password`* pelas credenciais do banco de dados. Para obter mais informações sobre como se conectar a instâncias de banco de dados MySQL e Aurora, consulte Conectar-se a uma instância de banco de dados executando o mecanismo de banco de dados MySQL e Conectar-se a um cluster de banco de dados do Athena no Guia

do usuário do Amazon RDS. `javax.jdo.option.ConnectionURL` é a string de conexão JDBC para um metastore JDBC. `javax.jdo.option.ConnectionDriverName` é o nome da classe de driver para um metastore JDBC.

Os drivers JDBC MySQL são instalados pelo Amazon EMR.

A propriedade "value" não pode conter espaços ou retornos de carro. Eles devem aparecer todos em uma única linha.

```
[
    {
      "Classification": "hive-site",
      "Properties": {
        "javax.jdo.option.ConnectionURL": "jdbc:mysql://hostname:3306/hive?
createDatabaseIfNotExist=true",
        "javax.jdo.option.ConnectionDriverName": "org.mariadb.jdbc.Driver",
        "javax.jdo.option.ConnectionUserName": "username",
        "javax.jdo.option.ConnectionPassword": "password"
      }
    }
  ]
```

b. Faça referência ao `hiveConfiguration.json` arquivo ao criar o cluster, conforme mostrado no AWS CLI comando a seguir. Neste comando, o arquivo é armazenado localmente, você também pode carregar o arquivo para o Amazon S3 e consultá-lo, por exemplo, `s3://DOC-EXAMPLE-BUCKET/hiveConfiguration.json`.

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge --
instance-count 2 \
--applications Name=Hive --configurations file://hiveConfiguration.json --use-
default-roles
```

4. Conecte-se ao nó primário do cluster.

Para obter informações sobre como se conectar ao nó primário, consulte Connect to the primary node using SSH no Guia de gerenciamento do Amazon EMR.

5. Crie tabelas do Hive especificando o local no Amazon S3 ao digitar um comando semelhante ao seguinte:

```
CREATE EXTERNAL TABLE IF NOT EXISTS table_name
(
key int,
value int
)
LOCATION s3://DOC-EXAMPLE-BUCKET/hdfs/
```

6. Adicione seu script do Hive ao cluster em execução.

O cluster do Hive é executado com o uso do metastore localizado no Amazon RDS. Execute todos os clusters Hive adicionais que compartilham essa metastore, especificando a localização da metastore.

Usar o driver JDBC do Hive

Você pode usar ferramentas populares de business intelligence, como Microsoft Excel, MicroStrategyQlikView, e Tableau, com o Amazon EMR para explorar e visualizar seus dados. Muitas dessas ferramentas exigem um driver JDBC (Java Database Connectivity) ou ODBC (Open Database Connectivity). O Amazon EMR é compatível com as conectividades do JDBC e do ODBC.

O exemplo abaixo demonstra o uso do SQL Workbench/J como um cliente SQL para conexão a um cluster do Hive no Amazon EMR. Para drivers adicionais, consulte Usar ferramentas de inteligência comercial com o Amazon EMR.

Antes de instalar e trabalhar com o SQL Workbench/J, faça download do pacote de driver e instale o driver. Os drivers incluídos no pacote são compatíveis com as versões do Hive disponíveis nas versões 4.0 e posteriores do Amazon EMR. Para obter notas de release e documentação detalhadas, consulte a documentação do PDF incluída no pacote.

- Download do pacote de drivers JDBC do Hive mais recente

  http://awssupportdatasvcs.com/bootstrap-actions/Simba/latest/
- Versões mais antigas do driver JDBC do Hive

http://awssupportdatasvcs.com/bootstrap-actions/Simba/

Para instalar e configurar o SQL Workbench

1. Faça o download do cliente do SQL Workbench/J para o seu sistema operacional de http://www.sql-workbench.net/downloads.html.
2. Instale o SQL Workbench/J. Para obter mais informações, consulte Instalar e iniciar o SQL Workbench/J (apenas em inglês) no SQL Workbench/J User's Manual.
3. Usuários do Linux, Unix, Mac OS X: em uma sessão de terminal, crie um túnel de SSH para o nó principal do cluster usando o comando a seguir. *master-public-dns-name*Substitua pelo nome DNS público do nó principal e *path-to-key-file* pelo local e nome do arquivo do seu arquivo de chave EC2 privada (`.pem`) da Amazon.

   ```
   ssh -o ServerAliveInterval=10 -i path-to-key-file -N -L 10000:localhost:10000
    hadoop@master-public-dns-name
   ```

   Usuários do Windows: em uma sessão do PuTTY, crie um túnel SSH para o nó principal do cluster (usando o encaminhamento de porta local) com `10000` para Porta de origem e *master-public-dns-name*`:10000` para Destino. Substitua *master-public-dns-name* pelo nome DNS público do nó principal.
4. Adicione o driver do JDBC ao SQL Workbench.

   a. Na caixa de diálogo Select Connection Profile (Selecionar perfil de conexão), clique em Manage Drivers (Gerenciar drivers).

   b. Clique no ícone Create a new entry (Criar uma nova entrada) (página em branco).

   c. No campo Name (Nome), digite **`Hive JDBC`**.

   d. Em Library (Biblioteca), clique no ícone Select the JAR file(s) (Selecionar os arquivos JAR).

   e. Navegue até o local que contém os drivers extraídos. Selecione os drivers incluídos na versão do pacote de drivers JDBC que você baixou e clique em Abrir.

      Por exemplo, seu pacote de drivers JDBC pode incluir o seguinte. JARs

      ```
      hive_metastore.jar
      hive_service.jar
      HiveJDBC41.jar
      libfb303-0.9.0.jar
      ```

```
libthrift-0.9.0.jar
log4j-1.2.14.jar
ql.jar
slf4j-api-1.5.11.jar
slf4j-log4j12-1.5.11.jar
TCLIServiceClient.jar
zookeeper-3.4.6.jar
```

f. Na caixa de diálogo Please select one driver (Selecione um driver), selecione `com.amazon.hive.jdbc41.HS2Driver`OK.

5. Quando você retornar para a caixa de diálogo Manage Drivers (Gerenciar drivers), verifique se o campo Classname (Nome da classe) está preenchido e selecione OK.
6. Quando você retornar para a caixa de diálogo Select Connection Profile (Selecionar perfil de conexão), verifique se o campo Driver está definido como Hive JDBC (JDBC do Hive) e forneça a string de conexão do JDBC no campo URL: `jdbc:hive2://localhost:10000/default`.
7. Selecione OK para conectar. Depois que a conexão estiver concluída, os detalhes da conexão irão aparecer na parte superior da janela do SQL Workbench/J.

Para obter mais informações sobre como usar o Hive e a interface JDBC, consulte a documentação do Hive JDBCInterface no HiveClientApache Hive.

Melhorar a performance do Hive

O Amazon EMR oferece atributos para ajudar na otimização da performance ao usar o Hive para consultar, ler e gravar dados salvos no Amazon S3.

O S3 Select pode melhorar a performance de consultas para arquivos CSV e JSON em algumas aplicações ao enviar o processamento para o Amazon S3.

O committer otimizado do EMRFS S3 é uma alternativa à OutputCommitterclasse, que elimina as operações de lista e renomeação para melhorar o desempenho ao gravar arquivos no Amazon S3 usando o EMRFS.

Tópicos

- Habilitar o confirmador otimizado para EMRFS S3 do Hive
- Usar o S3 Select com o Hive para melhorar a performance
- Otimização do MSCK

Habilitar o confirmador otimizado para EMRFS S3 do Hive

O confirmador otimizado para EMRFS S3 do Hive é uma forma alternativa pela qual o Hive no EMR grava arquivos para inserir consultas ao ser usado o EMRFS. O confirmador elimina as operações de listar e renomear feitas no Amazon S3 e melhora a performance da aplicação. O atributo está disponível desde o EMR 5.34 e o EMR 6.5.

Habilitar o confirmador

Se você quiser permitir que o Hive no EMR use `HiveEMRFSOptimizedCommitter` para confirmar dados como padrão para todas as tabelas externas e gerenciadas do Hive, use a seguinte configuração `hive-site` nos clusters do EMR 6.5.0 ou do EMR 5.34.0.

```
[
   {
      "classification": "hive-site",
      "properties": {
         "hive.blobstore.use.output-committer": "true"
      }
   }
]
```

> **Note**
>
> Não ative esse atributo quando `hive.exec.parallel` estiver definido como `true`.

Limitações

As restrições básicas a seguir se aplicam a tags:

- Não há suporte para habilitar o Hive para mesclar arquivos pequenos automaticamente. A lógica de confirmação padrão do Hive será usada mesmo quando o confirmador otimizado estiver habilitado.
- As tabelas ACID do Hive não são suportadas. A lógica de confirmação padrão do Hive será usada mesmo quando o confirmador otimizado estiver habilitado.
- A nomenclatura de nomeação para arquivos gravados foi alterada de `<task_id>_<attempt_id>_<copy_n>` para

`<task_id>_<attempt_id>_<copy_n>_<query_id>` do Hive. Por exemplo, um arquivo denominado

`s3://warehouse/table/partition=1/000000_0` será alterado para `s3://warehouse/table/partition=1/000000_0-hadoop_20210714130459_ba7c23ec-5695-4947-9d98-8a40ef759222-1`. A `query_id` aqui é uma combinação de nome de usuário, carimbo de data/hora e UUID.

- Quando partições personalizadas estão em sistemas de arquivos diferentes (HDFS, S3), esse atributo é automaticamente desabilitado. A lógica de confirmação padrão do Hive será usada quando estiver habilitada.

Usar o S3 Select com o Hive para melhorar a performance

> **Important**
>
> O Amazon S3 Select não está mais disponível para novos clientes. Os clientes existentes do Amazon S3 Select podem continuar usando o recurso normalmente. Saiba mais

Com as versões 5.18.0 e posteriores do Amazon EMR, você pode usar o S3 Select com o Hive no Amazon EMR. O S3 Select possibilita que as aplicações recuperem apenas um subconjunto dos dados de um objeto. Para o Amazon EMR, o trabalho de computação de filtrar grandes conjuntos de dados para processamento é enviado do cluster para o Amazon S3, o que pode melhorar a performance em algumas aplicações e reduz o volume de dados transferidos entre o Amazon EMR e o Amazon S3.

O S3 Select é compatível com tabelas do Hive com base em arquivos CSV e JSON e definindo a variável de configuração `s3select.filter` como `true` durante a sessão do Hive. Para ter mais informações e exemplos, consulte Especificar o S3 Select no código.

O S3 Select é adequado para minha aplicação?

Recomendamos que você avalie seus aplicativos com e sem o S3 Selecione para ver se o uso pode ser adequado para o aplicativo.

Use as seguintes diretrizes para determinar se o seu aplicativo é adequado para o uso do S3 Select:

- Sua consulta filtra mais de metade do conjunto de dados original.

- Os predicados do filtro de consulta usam colunas que têm um tipo de dados compatível com o Amazon S3 Select. Para obter mais informações, consulte Tipos de dados no Guia do usuário do Amazon Simple Storage Service.
- Sua conexão de rede entre o Amazon S3 e o cluster do Amazon EMR tem boa velocidade de transferência e largura de banda disponível. O Amazon S3 não compacta respostas HTTP. Portanto, é provável que o tamanho da resposta aumente para arquivos de entrada compactados.

Considerações e limitações

- A criptografia do lado do servidor do Amazon S3 com chaves de criptografia fornecidas pelo cliente (SSE-C) e a criptografia do lado do cliente não são compatíveis.
- A propriedade `AllowQuotedRecordDelimiters` não é compatível. Se essa propriedade for especificada, a consulta falhará.
- Somente arquivos CSV e JSON no formato UTF-8 são compatíveis. Não há suporte para várias linhas CSVs e JSON.
- Somente arquivos descompactados, gzip ou bzip2 são compatíveis.
- Caracteres de comentário na última linha não são compatíveis.
- Linhas vazias no final de um arquivo não são processadas.
- O Hive no Amazon EMR oferece suporte a tipos de dados primitivos compatíveis com o S3 Select. Para obter mais informações, consulte Tipos de dados no Guia do usuário do Amazon Simple Storage Service.

Especificar o S3 Select no código

Para usar o S3 Select em uma tabela do Hive, crie a tabela especificando `com.amazonaws.emr.s3select.hive.S3SelectableTextInputFormat` como o nome da classe `INPUTFORMAT` e especifique um valor para a propriedade `s3select.format` usando a cláusula `TBLPROPERTIES`.

Por padrão, o S3 Select está desabilitado quando você executa consultas. Habilite o S3 Select definindo `s3select.filter` como `true` em sua sessão do Hive, conforme mostrado a seguir. Os exemplos a seguir demonstram como especificar o S3 Select ao criar uma tabela do CSV subjacente e arquivos JSON e, em seguida, consultar a tabela com uma instrução simples do Select.

Example Instrução CREATE TABLE para tabela baseada em CSV

```
CREATE TABLE mys3selecttable (
col1 string,
col2 int,
col3 boolean
)
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
STORED AS
INPUTFORMAT
  'com.amazonaws.emr.s3select.hive.S3SelectableTextInputFormat'
OUTPUTFORMAT
  'org.apache.hadoop.hive.ql.io.HiveIgnoreKeyTextOutputFormat'
LOCATION 's3://path/to/mycsvfile/'
TBLPROPERTIES (
  "s3select.format" = "csv",
  "s3select.headerInfo" = "ignore"
);
```

Example Instrução CREATE TABLE para tabela baseada em JSON

```
CREATE TABLE mys3selecttable (
col1 string,
col2 int,
col3 boolean
)
ROW FORMAT SERDE 'org.apache.hive.hcatalog.data.JsonSerDe'
STORED AS
INPUTFORMAT
  'com.amazonaws.emr.s3select.hive.S3SelectableTextInputFormat'
OUTPUTFORMAT
  'org.apache.hadoop.hive.ql.io.HiveIgnoreKeyTextOutputFormat'
LOCATION 's3://path/to/json/'
TBLPROPERTIES (
  "s3select.format" = "json"
);
```

Example Instrução SELECT TABLE

```
SET s3select.filter=true;
SELECT * FROM mys3selecttable WHERE col2 > 10;
```

Otimização do MSCK

O Hive armazena uma lista de partições para cada tabela no seu metastore. No entanto, quando as partições são adicionadas ou removidas diretamente do sistema de arquivos, o metastore do Hive não toma conhecimento dessas alterações. O comando MSCK atualiza os metadados da partição no metastore do Hive para partições que foram adicionadas ou removidas diretamente do sistema de arquivos. A sintaxe do comando é:

```
MSCK [REPAIR] TABLE table_name [ADD/DROP/SYNC PARTITIONS];
```

O Hive implementa esse comando da seguinte forma:

1. O Hive recupera todas as partições da tabela do metastore. Em seguida, com base na lista de caminhos de partição que não existem no sistema de arquivos, ele cria uma lista de partições a serem descartadas do metastore.
2. O Hive reúne os caminhos de partição presentes no sistema de arquivos, compara-os com a lista de partições do metastore e gera uma lista de partições que precisam ser adicionadas ao metastore.
3. O Hive atualiza a metastore usando o modo `ADD`, `DROP` ou `SYNC`.

> **Note**
>
> Quando há muitas partições no metastore, a etapa para verificar se uma partição não existe no sistema de arquivos leva muito tempo para ser executada porque a chamada de API `exists` do sistema de arquivos deve ser feita para cada partição.

No Amazon EMR 6.5.0, o Hive introduziu um sinalizador denominado `hive.emr.optimize.msck.fs.check`. Quando habilitado, esse sinalizador faz com que o Hive verifique a presença de uma partição na lista de caminhos de partição do sistema de arquivos gerada na etapa 2 acima, em vez de fazer chamadas de API do sistema de arquivos. No Amazon EMR 6.8.0, o Hive habilitou essa otimização por padrão, eliminando a necessidade de definir o sinalizador `hive.emr.optimize.msck.fs.check`.

Usar o Live Long and Process (LLAP) do Hive

O Amazon EMR 6.0.0 é compatível com a funcionalidade Live Long and Process (LLAP) para o Hive. O LLAP usa daemons persistentes com cache inteligente na memória para melhorar o desempenho da consulta em comparação com o modo de execução do contêiner Tez padrão anterior.

Os daemons do LLAP do Hive são gerenciados e executados como um serviço do YARN. Como um serviço do YARN pode ser considerado um aplicativo do YARN de longa execução, alguns dos recursos de cluster são dedicados ao LLAP do Hive e não podem ser usados para outras cargas de trabalho. Para obter mais informações, consulte LLAP e YARN Service API.

Habilitar o LLAP do Hive no Amazon EMR

Para habilitar o LLAP do Hive no Amazon EMR, forneça a configuração a seguir ao iniciar um cluster.

```
[
  {
    "Classification": "hive",
    "Properties": {
      "hive.llap.enabled": "true"
    }
  }
]
```

Para obter mais informações, consulte Configurar aplicações.

Por padrão, o Amazon EMR aloca cerca de 60 por cento dos recursos do YARN do cluster para daemons do LLAP do Hive. É possível configurar a porcentagem de recurso do YARN do cluster alocado ao LLAP do Hive e o número de nós core e de tarefas a serem considerados para a alocação de LLAP do Hive.

Por exemplo, a configuração a seguir inicia o LLAP do Hive com três daemons em três nós core ou de tarefa e aloca 40% do recurso do YARN dos três nós core ou de tarefa para os daemons do LLAP do Hive.

```
[
  {
    "Classification": "hive",
    "Properties": {
      "hive.llap.enabled": "true",
      "hive.llap.percent-allocation": "0.4",
```

```
      "hive.llap.num-instances": "3"
    }
  }
]
```

É possível usar as configurações `hive-site` a seguir na API de classificação para substituir as configurações de recurso do LLAP padrão.

Propriedade	Descrição
hive.llap.daemon.yarn.container.mb	Tamanho total do contêiner do daemon do LLAP (em MB)
hive.llap.daemon.memory.per.instance.mb	A memória total usada pelos executores no contêiner do daemon do LLAP (em MB)
hive.llap.io.memory.size	Tamanho do cache de entrada/saída do LLAP
hive.llap.daemon.num.executors	Número de executores por daemon do LLAP

Iniciar manualmente o LLAP do Hive no cluster

Todas as dependências e configurações usadas pelo LLAP são empacotadas no arquivamento tar do LLAP como parte da inicialização do cluster. Se o LLAP estiver habilitado usando `"hive.llap.enabled": "true"`, recomendamos que você use a reconfiguração do Amazon EMR para fazer alterações de configuração no LLAP.

Caso contrário, para qualquer alteração manual no `hive-site.xml`, é necessário recriar o arquivamento tar do LLAP usando o comando `hive --service llap`, conforme demonstrado no exemplo a seguir.

```
# Define how many resources you want to allocate to Hive LLAP

LLAP_INSTANCES=<how many llap daemons to run on cluster>
LLAP_SIZE=<total container size per llap daemon>
LLAP_EXECUTORS=<number of executors per daemon>
LLAP_XMX=<Memory used by executors>
LLAP_CACHE=<Max cache size for IO allocator>
```

```
yarn app -enableFastLaunch

hive --service llap \
--instances $LLAP_INSTANCES \
--size ${LLAP_SIZE}m \
--executors $LLAP_EXECUTORS \
--xmx ${LLAP_XMX}m \
--cache ${LLAP_CACHE}m \
--name llap0 \
--auxhbase=false \
--startImmediately
```

Verificar o status do LLAP do Hive

Use o comando a seguir para verificar o status do LLAP do Hive pelo Hive.

```
hive --service llapstatus
```

Use o comando a seguir para verificar o status do LLAP do Hive usando o YARN.

```
yarn app -status (name-of-llap-service)

# example:
yarn app -status llap0 | jq
```

Iniciar ou interromper o LLAP do Hive

Como o LLAP do Hive é executado como um serviço persistente do YARN, você interrompe ou reinicia serviço do YARN para interromper ou reiniciar o LLAP do Hive. Os comandos a seguir demonstram isso.

```
yarn app -stop llap0
yarn app -start llap0
```

Redimensionar o número de daemons do LLAP do Hive

Use o comando a seguir para reduzir o número de instâncias do LLAP.

```
yarn app -flex llap0 -component llap -1
```

Para obter mais informações, consulte Flex a component of a service.

Criptografia no Hive

Esta seção descreve os tipos de criptografia compatíveis com o Amazon EMR.

Criptografia modular em Parquet no Hive

A criptografia modular em Parquet fornece controle de acesso e criptografia em nível colunar para aprimorar a privacidade e a integridade dos dados armazenados no formato de arquivo Parquet. Esse atributo está disponível no Hive no Amazon EMR desde a versão 6.6.0.

As soluções anteriormente compatíveis para segurança e integridade, que incluem a criptografia de arquivos ou a criptografia da camada de armazenamento, estão descritas em Opções de criptografia no Guia de gerenciamento do Amazon EMR. Essas soluções podem ser usadas para arquivos Parquet, mas o uso dos novos atributos do mecanismo de criptografia integrado do Parquet fornece acesso granular ao nível de coluna, além de melhorias na performance e na segurança. Saiba mais sobre esse atributo na página do Apache no github Criptografia modular em Parquet.

Os usuários transmitem configurações para leitores e gravadores do Parquet usando as configurações do Hadoop. As configurações detalhadas para que os usuários configurem leitores e gravadores para habilitar a criptografia e também alternar atributos avançados estão documentadas em PARQUET-1854: interface orientada por propriedades para o gerenciamento de criptografia em Parquet

Exemplos de uso

O exemplo a seguir aborda a criação e gravação em uma tabela do Hive usando AWS KMS para gerenciar chaves de criptografia.

1. Implemente um KmsClient para o AWS KMS serviço conforme descrito no documento PARQUET-1373: Ferramentas de gerenciamento de chaves de criptografia. O exemplo a seguir mostra um trecho de implementação.

```
package org.apache.parquet.crypto.keytools;

import com.amazonaws.AmazonClientException;
import com.amazonaws.AmazonServiceException;
import com.amazonaws.regions.Regions;
import com.amazonaws.services.kms.AWSKMS;
```

```
import com.amazonaws.services.kms.AWSKMSClientBuilder;
import com.amazonaws.services.kms.model.DecryptRequest;
import com.amazonaws.services.kms.model.EncryptRequest;
import com.amazonaws.util.Base64;
import org.apache.hadoop.conf.Configuration;
import org.apache.parquet.crypto.KeyAccessDeniedException;
import org.apache.parquet.crypto.ParquetCryptoRuntimeException;
import org.apache.parquet.crypto.keytools.KmsClient;
import org.slf4j.Logger;
import org.slf4j.LoggerFactory;

import java.nio.ByteBuffer;
import java.nio.charset.Charset;
import java.nio.charset.StandardCharsets;

public class AwsKmsClient implements KmsClient {

    private static final AWSKMS AWSKMS_CLIENT = AWSKMSClientBuilder
            .standard()
            .withRegion(Regions.US_WEST_2)
            .build();
    public static final Logger LOG = LoggerFactory.getLogger(AwsKmsClient.class);

    private String kmsToken;
    private Configuration hadoopConfiguration;

    @Override
    public void initialize(Configuration configuration, String kmsInstanceID, String
 kmsInstanceURL, String accessToken) throws KeyAccessDeniedException {
        hadoopConfiguration = configuration;
        kmsToken = accessToken;

    }

    @Override
    public String wrapKey(byte[] keyBytes, String masterKeyIdentifier) throws
 KeyAccessDeniedException {
        String value = null;
        try {
            ByteBuffer plaintext = ByteBuffer.wrap(keyBytes);

            EncryptRequest req = new
 EncryptRequest().withKeyId(masterKeyIdentifier).withPlaintext(plaintext);
            ByteBuffer ciphertext = AWSKMS_CLIENT.encrypt(req).getCiphertextBlob();
```

```
            byte[] base64EncodedValue = Base64.encode(ciphertext.array());
            value = new String(base64EncodedValue, Charset.forName("UTF-8"));
        } catch (AmazonClientException ae) {
            throw new KeyAccessDeniedException(ae.getMessage());
        }
        return value;
    }

    @Override
    public byte[] unwrapKey(String wrappedKey, String masterKeyIdentifier) throws
 KeyAccessDeniedException {
        byte[] arr = null;
        try {
            ByteBuffer ciphertext  =
 ByteBuffer.wrap(Base64.decode(wrappedKey.getBytes(StandardCharsets.UTF_8)));
            DecryptRequest request = new
 DecryptRequest().withKeyId(masterKeyIdentifier).withCiphertextBlob(ciphertext);
            ByteBuffer decipheredtext =
 AWSKMS_CLIENT.decrypt(request).getPlaintext();
            arr = new byte[decipheredtext.remaining()];
            decipheredtext.get(arr);
        } catch (AmazonClientException ae) {
            throw new KeyAccessDeniedException(ae.getMessage());
        }
        return arr;
    }
}
```

2. Crie suas chaves de AWS KMS criptografia para o rodapé e para as colunas com suas funções do IAM tendo acesso conforme descrito em Criação de chaves no Guia do AWS Key Management Service desenvolvedor. O perfil do IAM padrão é EMR_ECS_default.
3. Na aplicação do Hive em um cluster do Amazon EMR, adicione o cliente acima usando a instrução `ADD JAR`, conforme descrito na documentação de recursos do Apache Hive. Este é um exemplo de instrução:

   ```
   ADD JAR 's3://location-to-custom-jar';
   ```

   Um método alternativo é adicionar o JAR ao `auxlib` do Hive usando uma ação de bootstrap. Veja a seguir um exemplo de linha a ser adicionada à ação de boostrap:

```
aws s3 cp 's3://location-to-custom-jar' /usr/lib/hive/auxlib
```

4. Defina as seguintes configurações:

```
set
 parquet.crypto.factory.class=org.apache.parquet.crypto.keytools.PropertiesDrivenCryptoFacto
set
 parquet.encryption.kms.client.class=org.apache.parquet.crypto.keytools.AwsKmsClient;
```

5. Crie uma tabela do Hive com formato Parquet e especifique as AWS KMS chaves em SERDEPROPERTIES e insira alguns dados nela:

```
CREATE TABLE my_table(name STRING, credit_card STRING)
ROW FORMAT SERDE 'org.apache.hadoop.hive.ql.io.parquet.serde.ParquetHiveSerDe’
WITH SERDEPROPERTIES (
  'parquet.encryption.column.key’=<aws-kms-key-id-for-column-1>: credit_card’,
  'parquet.encryption.footer.key’='<aws-kms-key-id-for-footer>’)
STORED AS parquet
LOCATION “s3://<bucket/<warehouse-location>/my_table”;

INSERT INTO my_table SELECT
java_method ('org.apache.commons.lang.RandomStringUtils','randomAlphabetic',5) as
 name,
java_method ('org.apache.commons.lang.RandomStringUtils','randomAlphabetic',10) as
 credit_card
from (select 1) x lateral view posexplode(split(space(100),' ')) pe as i,x;

select * from my_table;
```

6. Verifique se, ao criar uma tabela externa no mesmo local sem acesso às AWS KMS chaves (por exemplo, acesso negado à função do IAM), você não consegue ler os dados.

```
CREATE EXTERNAL TABLE ext_table (name STRING, credit_card STRING)
ROW FORMAT SERDE 'org.apache.hadoop.hive.ql.io.parquet.serde.ParquetHiveSerDe’
STORED AS parquet
LOCATION “s3://<bucket>/<warehouse-location>/my_table”;

SELECT * FROM ext_table;
```

7. A última instrução deve gerar a seguinte exceção:

```
Failed with exception
 java.io.IOException:org.apache.parquet.crypto.KeyAccessDeniedException: Footer key:
 access denied
```

Criptografia em trânsito em 2 HiveServer

A partir da versão 6.9.0 do Amazon EMR, o HiveServer 2 (HS2) é habilitado para TLS/SSL como parte da configuração de segurança. Criptografia em trânsito em 2 HiveServer Isso afeta a forma como você se conecta à HS2 execução em um cluster do Amazon EMR com a criptografia em trânsito ativada. Para se conectar HS2, você deve modificar os valores dos `TRUSTSTORE_PASSWORD` parâmetros `TRUSTSTORE_PATH` e na URL do JDBC. O URL a seguir é um exemplo de uma conexão JDBC HS2 com os parâmetros necessários:

```
jdbc:hive2://HOST_NAME:10000/
default;ssl=true;sslTrustStore=TRUSTSTORE_PATH;trustStorePassword=TRUSTSTORE_PASSWORD
```

Use as instruções apropriadas para criptografia 2 dentro ou fora do cluster 2 abaixo HiveServer.

On-cluster HS2 access

Se você estiver acessando HiveServer 2 usando o cliente Beeline depois de fazer SSH no nó primário, faça referência `/etc/hadoop/conf/ssl-server.xml` para encontrar os valores dos `TRUSTSTORE_PASSWORD` parâmetros `TRUSTSTORE_PATH` e usando a configuração e. `ssl.server.truststore.location` `ssl.server.truststore.password`

Os exemplos de comandos a seguir podem ajudar você a recuperar essas configurações:

```
TRUSTSTORE_PATH=$(sed -n '/ssl.server.truststore.location/,+2p' /etc/hadoop/conf/
ssl-server.xml | awk -F "[><]" '/value/{print $3}')
TRUSTSTORE_PASSWORD=$(sed -n '/ssl.server.truststore.password/,+2p' /etc/hadoop/
conf/ssl-server.xml | awk -F "[><]" '/value/{print $3}')
```

Off-cluster HS2 access

Se você estiver acessando HiveServer 2 de um cliente fora do cluster do Amazon EMR, você pode usar uma das seguintes abordagens para obter o e: `TRUSTSTORE_PATH` `TRUSTSTORE_PASSWORD`

- Converta o arquivo PEM que foi criado durante a configuração de segurança em um arquivo JKS e use-o no URL de conexão JDBC. Por exemplo, com openssl e keytool, use os seguintes comandos:

```
openssl pkcs12 -export -in trustedCertificates.pem -inkey privateKey.pem -out
 trustedCertificates.p12 -name "certificate"
keytool -importkeystore -srckeystore trustedCertificates.p12 -srcstoretype pkcs12
 -destkeystore trustedCertificates.jks
```

- Como alternativa, faça referência a `/etc/hadoop/conf/ssl-server.xml` para encontrar os valores dos parâmetros `TRUSTSTORE_PATH` e `TRUSTSTORE_PASSWORD` usando a configuração `ssl.server.truststore.location` e `ssl.server.truststore.password`. Baixe o arquivo truststore para a máquina cliente e use o caminho na máquina cliente como `TRUSTSTORE_PATH`.

  Para obter mais informações sobre como acessar aplicações de um cliente fora do cluster do Amazon EMR, consulte Usar o driver JDBC do Hive.

Histórico de versões do Hive

A tabela a seguir informa a versão do Hive incluída em cada versão do Amazon EMR, além dos componentes instalados com a aplicação. Para obter as versões dos componentes em cada versão, consulte a seção Versão de componentes da sua versão em Versões 7.x do Amazon EMR, Versões de lançamento 6.x do Amazon EMR ou Versões de lançamento 5.x do Amazon EMR.

Informações de versão do Hive

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-7.9.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
		resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-7.8.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-7.7.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-7.6.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-7.5.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-7.4.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-7.3.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-7.2.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.36.2	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-7.1.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-7.0.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-6.15.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-6.14.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-6.13.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-6.12.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-6.11.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-6.11.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-6.10.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-6.10.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-6.9.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-6.9.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.8.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-6.8.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.7.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.36.1	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-5.36.0	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-6.6.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-5.35.0	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-6.5.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.4.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-6.3.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.3.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-6.2.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.2.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-6.1.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.1.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-6.0.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.0.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.34.0	2.3.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-5.33.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.33.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-5.32.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.32.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-5.31.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.31.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-5.30.2	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.30.1	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-5.30.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.29.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mysql-server, tez-on-yarn
emr-5.28.1	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.28.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mysql-server, tez-on-yarn
emr-5.27.1	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.27.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.26.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.25.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.24.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.24.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.23.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.23.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.22.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.21.2	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.21.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.21.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.20.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.20.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.19.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.19.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.18.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.18.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.17.2	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.17.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.17.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.16.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.16.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.15.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.15.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.14.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.14.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.14.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.13.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.13.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.12.3	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.12.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.12.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.12.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.11.4	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.11.3	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.11.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.11.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.11.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.10.1	2.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.10.0	2.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.9.1	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.9.0	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.8.3	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.8.2	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.8.1	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.8.0	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.7.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.7.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.6.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.6.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.5.4	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.5.3	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.5.2	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.5.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.5.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.4.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.4.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.3.2	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.3.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.3.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.2.3	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.2.2	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.2.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.2.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.1.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.1.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.0.3	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.0.2	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-5.0.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.0.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-4.9.6	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.9.5	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-4.9.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.9.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-4.9.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.9.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-4.8.5	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.8.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-4.8.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.8.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-4.8.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.8.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-4.7.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.7.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-4.7.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.7.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-4.7.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.6.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-4.6.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server
emr-4.5.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-4.4.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server
emr-4.3.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server

Rótulo de versão do Amazon EMR	Versão do Hive	Componentes instalados com o Hive
emr-4.2.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server
emr-4.1.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server
emr-4.0.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server

Notas da versão do Hive por versão

Tópicos

- Amazon EMR 7.8.0 - Notas de lançamento do Hive
- Amazon EMR 7.7.0 - Notas de lançamento do Hive
- Amazon EMR 7.6.0 - Notas de lançamento do Hive
- Amazon EMR 7.5.0 - Notas de lançamento do Hive
- Amazon EMR 7.4.0 — Notas de lançamento do Hive
- Amazon EMR 7.3.0 — Notas de lançamento do Hive
- Amazon EMR 7.2.0 — Notas de lançamento do Hive
- Amazon EMR 7.1.0: notas da versão do Hive
- Amazon EMR 7.0.0 - Notas de lançamento do Hive
- Amazon EMR 6.15.0: notas da versão do Hive
- Amazon EMR 6.14.0 - Notas da versão do Hive
- Amazon EMR 6.13.0 - Notas da versão do Hive
- Amazon EMR 6.12.0 - Notas da versão do Hive
- Amazon EMR 6.11.0 - Notas da versão do Hive
- Amazon EMR 6.10.0 - Notas da versão do Hive
- Amazon EMR 6.9.0 - Notas da versão do Hive
- Amazon EMR 6.8.0 - Notas da versão do Hive
- Amazon EMR 6.7.0 - Notas da versão do Hive
- Amazon EMR 6.6.0 - Notas da versão do Hive

Amazon EMR 7.8.0 - Notas de lançamento do Hive

Amazon EMR 7.8.0 - Alterações no Hive

Tipo	Descrição
Correção de bugs	Corrige CVE-2024-23953: Apache Hive: Timing Attack Against Signature no LLAP.

Problemas conhecidos

- Para consultas de sobregravação do Hive Insert com o Amazon S3 Express One Zone como local de saída, defina a configuração do site principal: como. `fs.s3a.directory.operations.purge.uploads` `false`

Amazon EMR 7.7.0 - Notas de lançamento do Hive

Amazon EMR 7.7.0 - Alterações no Hive

Tipo	Descrição
Correção de bugs	Corrige CVE-2024-29869: Apache Hive: arquivo de credenciais criado com permissões não restritivas.
Correção de bugs	Corrige SemanticException quando uma política de filtragem em nível de linha é habilitada no Apache Ran.
Correção de bugs	Desative o Tez Async Init RR quando o LLAP ou o ACID estiverem habilitados.

Problemas conhecidos

- Para consultas de sobregravação do Hive Insert com o Amazon S3 Express One Zone como local de saída, defina a configuração do site principal: como. `fs.s3a.directory.operations.purge.uploads` `false`

Amazon EMR 7.6.0 - Notas de lançamento do Hive

Amazon EMR 7.6.0 - Alterações no Hive

Tipo	Descrição
Melhoria	Adicionado recurso rápido de listagem de prefixos S3 para tabelas particionadas ORC não ACID
Recurso	Adicione suporte para Magic Committers for Hive Write Queries no sistema S3 AFile

Problemas conhecidos

- Para consultas de sobregravação do Hive Insert com o Amazon S3 Express One Zone como local de saída, defina a configuração do site principal: como. `fs.s3a.directory.operations.purge.uploads` `false`

Amazon EMR 7.6.0 - Novas configurações

Classificação	Name	Padrão	Descrição
hive-site	`hive.exec.fast.s3.partition.discovery.enabled`	true	Se deve usar a descoberta rápida de partições do S3 para o cálculo da divisão. Isso habilitará a listagem baseada em prefixos para formatos de arquivo compatíve is: ORC. Observe que esse recurso usa um parâmetro da API S3 que a classe de armazenamento S3 Express One Zone

Classificação	Name	Padrão	Descrição
			não suporta. Ao usá-los, desative esse recurso.
hive-site	`hive.exec.fast.s3.partition.discovery.max.thread.threshold`	128	O grau máximo de paralelismo para descoberta rápida de partições S3.
hive-site	`hive.exec.fast.s3.partition.discovery.parallelism`	10	O grau de paralelismo de uma única execução de descoberta rápida de partições do S3. Essa configuração só tem efeito se `hive.exec.fast.s3.partition.discovery.enabled` estiver definida como `true`

Classificação	Name	Padrão	Descrição
hive-site	`hive.blobstore.output-committer.magic.track.commits.in.memory.enabled`	true	Sinalizar para ativar o Committer do Magic com o Hive deve rastrear todas as confirmações pendentes na memória? O committer Magic tem a opção de armazenar os dados de confirmação na memória, o que pode acelerar a TaskCommit operação fazendo menos chamadas S3. Essa configuração substitui a configuração do Hadoop `fs.s3a.committer.magic.track.commits.in.memory.enabled`

Classificação	Name	Padrão	Descrição
hive-site	`hive.blobstore.output-committer.dp.skip.task.staging.dir.creation`	true	Sinalize para ativar se o Magic Committer deve criar os caminhos de preparação de dp na blobstore? Esse sinalizador é aplicável somente ao rastrear confirmações na memória quando o Hive usa o Magic Committer via. `hive.blobstore.output-committer.magic.track.commits.in.memory.enabled` Por padrão, ele é definido como verdadeiro, mas só entra em vigor se `hive.blobstore.output-committer.magic.track.commits.in.memory.enabled` estiver ativado e salvar chamadas adicionais do S3 ou criar caminhos de tentativa

Classificação	Name	Padrão	Descrição
			de tarefa no blobstore .
hive-site	`hive.blobstore.output-committer.magic.disable.fs.cache.for.llap`	true	Sinalize para ativar se os caches FS do blobstore devem ser desativados nos fluxos de gravação para LLAP ao usar o Magic Committer . Esse sinalizador aparece quando o LLAP está ativado e, por padrão, é definido como verdadeiro.

Amazon EMR 7.5.0 - Notas de lançamento do Hive

Amazon EMR 7.5.0 - Alterações no Hive

Tipo	Descrição
Melhoria	Aumento do tempo máximo de espera pela abertura da sessão Tez ao tentar usar a sessão existente no HiveCLI para 10s
Melhoria	Configurações ajustadas para melhorar o desempenho em consultas de seleção simples com LIMIT

Amazon EMR 7.4.0 — Notas de lançamento do Hive

Amazon EMR 7.4.0 - Alterações no Hive

Tipo	Descrição
Upgrade	HIVE-28191: Atualize a versão do Hadoop para 3.4.0
Upgrade	Atualize o protobuf compartilhado do hadoop para 3.21
Upgrade	Atualize commons-cli para 1.5.0
Upgrade	Atualize commons-compress para 1.24.0
Upgrade	Atualize commons-io para 2.14.0
Upgrade	Atualize commons-lang3 para 3.21.0
Melhoria	Altere o tempo de espera até que a sessão Tez seja aberta ao tentar usar a sessão existente no HiveCLI para 10s
Melhoria	Ative o mecanismo de curto-circuito no Tez DAG para consultas de seleção simples com LIMIT
Melhoria	HIVE-21100: Permitir o nivelamento dos subdiretórios da tabela resultante do uso do mecanismo TEZ e da cláusula UNION
Correção de bugs	HIVE-25095: O comando Beeline/hive -e não consegue lidar com a consulta com aspas finais
Correção de bugs	HIVE-13781: O Tez Job falhou quando o diretório de partição não existe FileNotFo undException

Tipo	Descrição
Correção de bugs	HIVE-28480: Desative o SMB na incompatibilidade do gerador de hash de partição entre ramificações de junção no RS anterior

Amazon EMR 7.4.0 - Novas configurações

Classificação	Name	Padrão	Descrição
hive-site	hive.ignore.failure.partition.dir.not.found	false	Ignora a falha se a partição da tabela existir, mas o caminho real de armazenamento do objeto não existir.
hive-site	subdiretórios hive.tez.union.flatten.subdiretórios	false	Ao gravar dados em uma tabela e UNION ALL ser a última etapa da consulta, o Hive on Tez cria um subdiretório para cada ramificação do UNION ALL. Quando essa propriedade é ativada, os subdiretórios são removidos e os arquivos são renomeados e movidos para o diretório principal. Observe que isso não tem efeito quando o hive.blobstore.use

Classificação	Name	Padrão	Descrição
			.output-committer está ativado.

Amazon EMR 7.3.0 — Notas de lançamento do Hive

Amazon EMR 7.3.0 - Alterações no Hive

Tipo	Descrição
Recurso	HIVE-18728 — Web segura com SSL. HCat
Melhoria	Support a configuração de credenciais de armazenamento de chaves SSL para a interface web do daemon LLAP.
Melhoria	Forneça a opção de controlar a verificação do nome de host SSL para o servidor Hive Metastore.
Correção de bugs	HIVE-26541 — NPE ao iniciar o serviço Web. HCat
Correção de bugs	HIVE-23011 — O otimizador de trabalho compartilhado deve verificar os predicados residuais ao comparar as junções.
Correção de bugs	Corrija javax.security.sasl. SaslException: Não há uma camada de proteção comum entre o cliente e o servidor entre o HMS e o Namenode quando a criptografia em trânsito está ativada.
Correção de bugs	Corrija IOExceptiononde o final da divisão orc se sobrepõe ao início da localização de um bloco.
Correção de bugs	Use o delimitador do nome da coluna em vez de sempre dividir por vírgula quando os nomes

Tipo	Descrição
	das colunas contiverem caracteres de vírgula e usar. CSVSerde

Amazon EMR 7.3.0 - Novas configurações

Classificação	Name	Padrão	Descrição
hcatalog-webhcat-site	templeton.use.ssl	false	Defina isso como verdadeiro para usar a criptografia SSL para o HCat servidor Web.
hcatalog-webhcat-site	templeton.keystore.path		Local de armazenamento de chaves do certificado SSL para servidor WebHCat .
hcatalog-webhcat-site	templeton.keystore.senha		Senha do repositório de chaves do certificado SSL para servidor WebHCat .
hcatalog-webhcat-site	templeton.ssl.protocol.blacklist	SSLv2, SSLv3	Versões SSL a serem desativadas para o HCat servidor Web.
hcatalog-webhcat-site	templeton.host	0.0.0.0	O endereço do host que o HCat servidor Web escutará.
hive-site	hive.metastore.ssl.enable.hostname.verification	false	Controle a verificação do nome do host durante o handshake SSL/TLS.

Classificação	Name	Padrão	Descrição
hive-site	hive.llap.daemon.web.ssl.keystore.path		Localização do repositório de chaves do certificado SSL para a interface web do daemon LLAP.
hive-site	hive.llap.daemon.web.ssl.keystore.password		Senha do repositório de chaves do certificado SSL para a interface web do daemon LLAP.
hive-site	hive.metastore.hadoop.rpc.protection.override.to.authentication	false	Quando ativado, o HMS sempre substitui o valor de hadoop.rpc.protection para autenticação em seu conjunto de configurações.

Amazon EMR 7.2.0 — Notas de lançamento do Hive

Amazon EMR 7.2.0 - Mudanças no Hive

Tipo	Descrição
Upgrade	Parquet 1.13.1 — O Parquet foi atualizado para 113.1.
Melhoria	HIVE-12930 — Support SSL shuffle para LLAP.
Melhoria	HIVE-23062 — Hive para verificar o URL do Yarn RM no modo TLS e Yarn HA para uma fila Tez personalizada.

Tipo	Descrição
Correção de bugs	HIVE-27952 — O Hive não consegue criar SslContextFactory quando tem vários certifica dos. KeyStore
Correção de bugs	HIVE-28085 — YarnQueueHelper falha ao acessar o YARN habilitado para HTTPS. WebService
Correção de bugs	HIVE-26436 — Hive on MR NullPointerExcepti on quando InitializeOp não foi chamado e fechou a chamada. Se o operador não tiver sido inicializado, pule o fechamento..

Amazon EMR 7.2.0 - Novas configurações

Classificação	Name	Padrão	Descrição
hive-site	hive.llap.shuffle. ssl.enabled	false	Defina como true junto com tez.runti me.shuffle.ssl.ena ble para ativar o SSL shuffle para LLAP.

Amazon EMR 7.1.0: notas da versão do Hive

Amazon EMR 7.1.0: alterações no Hive

Tipo	Descrição
Correção de bugs	HIVE-24381: a entrada de texto compactado retorna 0 linhas se ignorar cabeçalho ou rodapé estiver incluído.

Tipo	Descrição
Correção de bugs	HIVE-24190 — LLAP: ShuffleHandler pode retornar DISK_ERROR_EXCEPTION de acordo com TEZ-4233.
Correção de bugs	HIVE-23073: sombrear Netty.
Correção de bugs	HIVE-23073: sombrear Netty e atualizar para Netty 4.1.48.Final.
Correção de bugs	HIVE-23148: o fluxo do cliente externo do LLAP é interrompido devido ao sombreamento do Netty.
Correção de bugs	HIVE-25180: atualiza o Netty.
Correção de bugs	HIVE-24524 — LLAP ShuffleHandler: atualize para o Netty4 e remova a dependência do Netty3 do hive sempre que possível.
Correção de bugs	HIVE-28000: Hive QL: a cláusula "not in" fornece resultados incorretos quando a coerção de tipo não pode ocorrer.
Correção de bugs	HIVE-27993 — O Netty4 ShuffleHandler deve usar 1 fio principal.
Upgrade	Atualiza o Netty para 4.1.100.Final
Upgrade	Atualiza o Jetty para 9.4.53.v20231009
Upgrade	Atualiza o Zookeeper para 3.9.1

Amazon EMR 7.1.0: alterações no Hive

- O Amazon EMR 7.1 atualiza o Hive para o Netty 4.1.100.Final para resolver as vulnerabilidades de segurança no Netty3. Como depende do netty3, o Hive não hive-druid-handler tem o `hive-druid-handler` JAR no classpath do Hive no Amazon EMR 7.1. Uma próxima versão do

Amazon EMR o incluirá no classpath do Hive assim que o manipulador Druid oferecer suporte à versão 4.1.100.Final ou posterior do Netty. Entre em contato com o AWS suporte se precisar do `hive-druid-handler` JAR nas versões 7.1 ou superiores do Amazon EMR.

Amazon EMR 7.0.0 - Notas de lançamento do Hive

Amazon EMR 7.0.0 - Alterações no Hive

Tipo	Descrição
Upgrade	O Hive Runtime agora usa o Java 17 por padrão. Consulte o Guia de versão do EMR 7.0.0 para obter mais detalhes.
Backport	HIVE-17709: remova as referências do Sun.Misc.Cleaner
Correção de bugs	Desative o Tez Async Init RR quando o LLAP ou o ACID estiverem habilitados

Amazon EMR 6.15.0: notas da versão do Hive

Amazon EMR 6.15.0: alterações no Hive

Tipo	Descrição
Recurso	Suporte para TEZ-4397: para a abertura de divisão assíncrona do Tez, o Hive agora oferece suporte às configurações do Tez descritas em Abertura de divisão assíncrona do Tez.
Correção de bugs	HIVE-25400: move a atualização do deslocamento em `BytesColumnVector` para `setValPreallocated` .

Tipo	Descrição
Correção de bugs	HIVE-25190: corrige muitas pequenas alocações em `BytesColumnVector` .
Correção de bugs	Empacotar módulos netty com o servidor llap para evitar NoClassDefFoundexceções ao iniciar o LLapDaemon em nós de trabalho.
Upgrade	Atualização do Apache Hadoop para 3.3.6.
Upgrade	HIVE-26684: atualiza `maven-shade-plugin` para 3.4.1.
Melhoria	Para reduzir o tempo de inicialização do cluster do Amazon EMR, remova 15 segundos do tempo de espera do script de inicialização HCatalog.

Amazon EMR 6.14.0 - Notas da versão do Hive

Amazon EMR 6.14.0 - Alterações no Hive

Tipo	Descrição
Melhoria	HIVE-26762: remova a remoção do operando em `HiveFilterSetOpTransposeRule`
Correção de bugs	HIVE-27582: Não armazene em cache HBase o formato de entrada da tabela em FetchOperator
Correção de bugs	HIVE-26452: NPE ao converter JOIN para MAPJOIN uma coluna referenciada mais de uma vez JOIN

Tipo	Descrição
Correção de bugs	HIVE-26416: `AcidUtils.isRawFormatFile()` lança `InvalidProtocolBufferException` para o arquivo non-ORC
Correção de bugs	HIVE-26105: mostrar colunas mostra valores extras se a coluna Comentários contém caracteres chineses específicos
Correção de bugs	HIVE-25864: a otimização de consultas do Hive cria um plano errado para envio de predicados com a função de janela
Correção de bugs	HIVE-25224: várias declarações INSERT envolvendo tabelas com resultados diferentes em erro `bucketing_versions`
Correção de bugs	HIVE-24151: `MultiDelimitSerDe` muda os dados se as strings contém caracteres non-ASCII
Correção de bugs	HIVE-23606: (LLAP) atraso na limpeza de `DirectByteBuffer` para `EncodedReaderImpl`
Correção de bugs	HIVE-22165: sincronização introduzida pelo HIVE-14296 no `SessionManager.closeSession` causa alta latência em um servidor hive ocupado
Correção de bugs	HIVE-21304: tornar o uso da versão de compartimento mais robusto

Amazon EMR 6.13.0 - Notas da versão do Hive

Amazon EMR 6.13.0 - Alterações no Hive

Tipo	Descrição
Melhoria	Atualizar os scripts do Python para oferecer suporte ao Python3
Melhoria	HIVE-27097: Melhore a estratégia de repetição para cliente e servidor MetaStore
Correção de bugs	HIVE-21778: CBO: “Struct não é nulo” é avaliado como anulável, sempre causando ausência de filtro na consulta
Correção de bugs	HIVE-21009: adicionar a capacidade de o usuário definir o usuário de vinculação
Correção de bugs	HIVE-22661: a compactação falha em uma tabela sem bucket com dados carregados no caminho
Correção de bugs	HIVE-19718: adicionar partições em massa também busca a tabela para cada partição
Correção de bugs	HIVE-22173: consulta com várias vistas laterais trava durante a compilação
Correção de bugs	HIVE-27088: resultados incorretos quando junções internas e externas com filtros de pós-junção são mescladas
Correção de bugs	HIVE-21935: vetorização do Hive: performance degradada com UDF vetorizada
Correção de bugs	HIVE-25299: a conversão do carimbo de data e hora para tipos de dados numéricos está incorreta para fusos horários não UTC

Tipo	Descrição
Correção de bugs	HIVE-24626: LLAP: os threads do leitor podem ficar escassos se todos os threads de elevador de E/S estiverem ocupados para enfileirar para outros leitores com a fila cheia
Correção de bugs	HIVE-27029: falha na consulta do hive com erro de fechamento do sistema de arquivos, retrabalho feito para HIVE-26352
Correção de bugs	HIVE-26352: a verificação de acesso à fila do Tez falha com a exceção GSS na compactação
Correção de bugs	HIVE-24590: o registro em log de operações ainda vaza anexos do log4j
Correção de bugs	HIVE-24552: Possíveis conexões HMS vazam ou se acumulam em loadDynamicPartitions
Correção de bugs	HIVE-27069: resultados incorretos com junção de mapa de bucket
Correção de bugs	HIVE-27344: Adicione uma verificação nula em #close RecordReaderImpl
Correção de bugs	HIVE-27439: dar suporte a espaço em decimal
Correção de bugs	HIVE-27267: resultados incorretos ao ser feita a junção do mapa de buckets em uma coluna decimal com buckets com subconsulta
Correção de bugs	HIVE-21986: UI HiveServer da Web: definindo o cabeçalho de resposta Strict-Transport-S ecurity padrão
Correção de bugs	HIVE-22148: os tokens de delegação do S3A não são adicionados na configuração do trabalho do Compactador.

Tipo	Descrição
Correção de bugs	HIVE-22622: o Hive permite a criação de um struct com nomes de atributos duplicados
Correção de bugs	HIVE-22008: o operador LIKE deve correspon der uma entrada de várias linhas
Correção de bugs	HIVE-23144: LLAP: deixe a limpeza no ServiceStop QueryTracker
Correção de bugs	HIVE-22391: NPE na verificação do cache de resultados da consulta do Hive
Correção de bugs	HIVE-23305: NullPointerException em LlapTaskSchedulerService addNode devido à condição de corrida
Correção de bugs	HIVE-22178: Parquet lança depois FilterPre dicate CastException SchemaEvolution
Correção de bugs	HIVE-21517: Correção AggregateStatsCache
Correção de bugs	HIVE-21825: melhorar a mensagem de erro do cliente quando HA ativa/passiva estiver habilitada
Correção de bugs	HIVE-23389: pode levar a FilterMergeRule AssertionError
Correção de bugs	HIVE-22767: Beeline não analisa ponto e vírgula nos comentários corretamente
Correção de bugs	HIVE-22996: a BasicStats análise deve verificar proativamente se há uma string nula ou vazia
Correção de bugs	HIVE-22808: não suporta HiveRelFieldTrimmer HiveTableFunctionScan

Tipo	Descrição
Correção de bugs	HIVE-22437: cache de NPE de metadados LLAP no bloqueio de metadados.
Correção de bugs	HIVE-22606: AvroSerde registra avro.schema.literal no nível INFO
Correção de bugs	HIVE-22713: a propagação constante não deve ser feita para a estrutura Join-Fil(*)-RS
Correção de bugs	HIVE-21624: LLAP: as métricas da Cpu em nível de thread estão corrompidas
Correção de bugs	HIVE-22815: reduza a criação desnecessária de objetos do sistema de arquivos no MROutput
Correção de bugs	HIVE-23060: consulta com o erro “Grouping sets expression is not in GROUP BY key. Error encountered near token”
Correção de bugs	HIVE-22236: falha ao criar visualização ao selecionar Visualização contendo subconsulta `NOT IN`
Correção de bugs	HIVE-19886: os logs poderão ser direcionados para dois arquivos se —hiveconf hive.log.file for usado
Correção de bugs	HIVE-20620: colisões evidentes na inserção em tabelas MM classificadas com buckets com particionamento dinâmico
Correção de bugs	HIVE-14557: Nullpointer quando ambos e Mapjoin estão habilitados SkewJoin
Correção de bugs	HIVE-20471: problemas na obtenção do caminho padrão do banco de dados

Tipo	Descrição
Correção de bugs	HIVE-20598: Corrigir erros de digitação nos cálculos HiveAlgorithmsUtil
Correção de bugs	HIVE-14737: problema ao acessar /logs em uma interface de usuário da Web Kerberized Hive Server 2
Correção de bugs	HIVE-20733: O UDFOPEqual NS genérico não pode usar = nas descrições do plano
Correção de bugs	HIVE-20848: Depois de definir a UpdateInputAccessTimeHook consulta, falha com a tabela não encontrada.
Correção de bugs	HIVE-18929: O método humanReadableInt em HiveStringUtils .java tem uma condição de corrida.
Correção de bugs	HIVE-20841: LLAP: tornar as portas dinâmicas configuráveis
Correção de bugs	HIVE-20930: VectorCoalesce no modo FILTRO não tem efeito
Correção de bugs	HIVE-21007: semijunção e união podem levar a planos errados
Correção de bugs	HIVE-21074: a remoção de consultas de tabelas com buckets do Hive não funciona na condição IS NOT NULL
Correção de bugs	HIVE-21223: CachedStore retorna partição nula quando a partição não existe
Correção de bugs	HIVE-19625: NPE potencial e exceção real oculta em Hive#copyFiles

Tipo	Descrição
Correção de bugs	HIVE-17020: a eliminação de duplicação agressiva no RS pode remover incorretamente uma ramificação da árvore OP
Correção de bugs	HIVE-20168: Registro oculto ReduceSin kOperator
Correção de bugs	HIVE-20879: Usar null em uma expressão de projeção leva a CastException
Correção de bugs	HIVE-20888: TxnHandler: sort () chamado em listas imutáveis
Correção de bugs	HIVE-19948: não HiveCli está dividindo o comando por ponto e vírgula corretamente se as aspas estiverem dentro da string
Correção de bugs	HIVE-20621: GetOperationStatus chamado em resultset.next causando lentidão incremental
Correção de bugs	HIVE-20854: padrões sensatos: o intervalo de pulsação do Zookeeper do Hive é de 20 minutos, mudar para 2
Correção de bugs	HIVE-20330: O HCat carregador não pode manipular vários InputJobInfo objetos para um trabalho com várias entradas
Correção de bugs	HIVE-20787: A capa MapJoinBytesTableC ontainer DummyRow não suporta reutilização
Correção de bugs	HIVE-20331: a consulta com union all, vista lateral e junção falha com “cannot find parent in the child operator”
Correção de bugs	HIVE-19968: a exceção de UDF não é descartada

Tipo	Descrição
Correção de bugs	HIVE-20410: a substituição de inserção abortada na tabela transacional causa o erro “Not enough history available for…”
Correção de bugs	HIVE-20059: o streaming do Hive deve tentar o prefixo de sombra incondicionalmente na exceção
Correção de bugs	HIVE-19424: Entrada NPE MetaDataF ormatters
Correção de bugs	HIVE-20355: Limpe o parâmetro de.setSchema HiveConnection
Correção de bugs	HIVE-20858: O serializador não foi inicializado corretamente com a configuração em Utilitários. createEmptyBuckets
Correção de bugs	HIVE-20424: a ferramenta de esquema não deve poluir a história do beeline
Correção de bugs	HIVE-20338: LLAP: forçar file-id sintético para sistemas de arquivos que têm impls de protocolo HDFS com semântica de mutação POSIX
Correção de bugs	HIVE-11708: Operadores lógicos aumentam com NULL ClassCastExceptions
Correção de bugs	HIVE-21082: em HPL/SQL, a instrução declare não é compatível com a variável de digitar caractere
Correção de bugs	HIVE-16690: configurar a borda do produto cartesiano do Tez com base no tamanho do cluster do LLAP

Tipo	Descrição
Correção de bugs	HIVE-14516:. OrcInputFormat SplitGenerator.Chamada interna
Correção de bugs	HIVE-20981: transmissão/vazamentos AbstractRecordWriter HeapMemoryMonitor
Correção de bugs	HIVE-20043: HiveServer 2: SessionState tem um bloco de sincronização estática em torno de um AtomicBoolean
Correção de bugs	HIVE-20191: o aplicativo de PreCommit patch não falha se o patch estiver vazio
Correção de bugs	HIVE-20400: a ação de criar tabela deve sempre usar um caminho totalmente qualificado para evitar possíveis ambiguidades em FS
Correção de bugs	Adicionar uma verificação nula para skewedInfo antes de acessar colunas distorcidas

Amazon EMR 6.12.0 - Notas da versão do Hive

Amazon EMR 6.12.0 - Alterações no Hive

Tipo	Descrição
Melhoria	Adicionado suporte para runtime do JDK 11 e do JDK 17
Melhoria	Adicione suporte para consultar nomes de colunas de palavras-chave reservadas e sensíveis a maiúsculas e minúsculas quando usar o S3 Select. Para usar isso, defina a propriedade da tabela no formato “s3select.column.mapping” = "*column1*:*fieldName1* , *column2*:*fieldName2* ,..."

Tipo	Descrição
Melhoria	HIVE-23133: as operações numéricas podem ter resultados diferentes nas arquiteturas de hardware
Melhoria	HIVE-27145: Use StrictMath para funções matemáticas restantes como acompanhamento do HIVE-23133
Correção de bugs	Corrige a incompatibilidade de caracteres curinga nos HMS get_partitions_by_filter e get_num_partitions_by_filter causada pela portabilidade do HIVE-22900 no EMR Hive 6.4.0 APIs
Correção de bugs	HIVE-26736: falha na autorização para visualizações aninhadas e que têm a cláusula WITH
Correção de bugs	HIVE-22416: ausência dos logs de operação relacionados ao MR quando a execução paralela está habilitada
Correção de bugs	HIVE-19653: pushdown de predicados incorretos para groupby com conjuntos de agrupamento
Correção de bugs	HIVE-22094: falhas nas consultas com: hive.ql.exec.vector. ClassCastException DecimalColumnVectornão pode ser convertido em hive.ql.exec.vector.Decimal64 ColumnVec tor
Correção de bugs	HIVE-26340: o operador PTF vetorizado apresentará falha se a consulta tiver a função de janela em maiúsculas

Tipo	Descrição
Correção de bugs	HIVE-26184: COLLECT_SET com GROUP BY é muito lento quando algumas chaves estão muito distorcidas
Correção de bugs	HIVE-26373: ClassCastException ao ler carimbos de data/hora da tabela com dados Avro HBase
Correção de bugs	HIVE-26388: ClassCastException quando não há coluna do tipo string na tabela de origem da consulta CTAS Atualize HIVE-26172: Hive - Atualize o Ant para 1.10.11 devido ao CVE-2021-36373 e CVE-2021-36374
Correção de bugs	HIVE-26114: corrigir a conexão jdbc hiveserver2 usando o comando dfs com espaço de prefixo causará exceção
Correção de bugs	HIVE-26396: a função trunc tem um problema com a interceptação de precisão e o resultado tem muitos 0
Correção de bugs	HIVE-26446: HiveProtoLoggingHook falha ao preencher o campo para tabelas particionadas. TablesWritten
Correção de bugs	HIVE-26639: ConstantVectorExpression e não ExplainTask deve depender do conjunto de caracteres padrão
Correção de bugs	HIVE-22670: ArrayIndexOutOfBoundsException quando o leitor vetorizado é usado para ler um arquivo de parquet
Correção de bugs	HIVE-23607: problema de permissão: a criação de uma visão sobre outra visão é bem-sucedida, mas a visão alternativa apresenta falha

Tipo	Descrição
Correção de bugs	HIVE-25498: consulta com mais de 31 funções distintas de contagem retorna um resultado errado
Correção de bugs	HIVE-25780: DistinctExpansion cria mais de 64 conjuntos de agrupamento II
Correção de bugs	HIVE-23868: especificação da função de janela: suporte 0 anterior/seguinte
Correção de bugs	HIVE-24539: a geração do OrcInputFormat esquema deve respeitar o delimitador de coluna
Correção de bugs	HIVE-23476: LLAP: pré-alocar arenas para casos de mmap também
Correção de bugs	HIVE-25806: Possível vazamento em LlapCacheAwareFs - Parquet, LLAP IO
Correção de bugs	HIVE-23498: Desative o método HTTP Trace em ThriftHttpCliService
Correção de bugs	HIVE-25729: ThriftUnionObjectInspector deve ser notificado quando totalmente iniciado
Correção de bugs	HIVE-23846: evitar serialização e desserialização desnecessárias de bitvetores
Correção de bugs	HIVE-24233: a subconsulta except lança um nullpointer com cbo desativado
Correção de bugs	HIVE-24276: Vulnerabilidade de script cruzado (XSS) de HiveServer 2 loggerconf jsp
Correção de bugs	HIVE-25721: o resultado da junção externa está errado

Tipo	Descrição
Correção de bugs	HIVE-25223: selecionar com limite não retorna linhas na tabela não nativa
Correção de bugs	HIVE-25794: CombineHiveRecordReader: declarações de registro em um loop levam à pressão da memória
Correção de bugs	HIVE-23602: usar o pacote simultâneo do Java para o conjunto do processador de operações
Correção de bugs	HIVE-24045: nenhum registro em log relaciona do a quando o banco de dados padrão é criado
Correção de bugs	HIVE-24305: o esquema decimal avro não preenche adequadamente a escala/precisão se o valor estiver entre aspas
Correção de bugs	HIVE-25040: a cascata de banco de dados drop não pode remover funções persistentes
Correção de bugs	HIVE-23501: AOOB em VectorDeserializeRow quando tipos complexos são convertidos em tipos primitivos
Correção de bugs	HIVE-23704: o servidor HTTP do Thrift não manipula o identificador de autenticação corretamente
Correção de bugs	HIVE-23529: o CTAS danificado para uniontype quando row_deserialize
Correção de bugs	HIVE-24144: A getIdentifierQuote string in HiveDatabaseMetaData retorna um valor incorreto

Tipo	Descrição
Correção de bugs	HIVE-23850: permitir PPD quando o assunto não é uma coluna com conjuntos de agrupamento presentes
Correção de bugs	HIVE-25919: ClassCastException ao inserir o predicado da coluna booleana HBase StorageHandler
Correção de bugs	HIVE-25261: A nova tentativa HMSHandle r deve incluir uma breve descrição do MetaException alvo
Correção de bugs	HIVE-24792: potencial vazamento de thread em Operação
Correção de bugs	HIVE-23409: Se a reabertura do TezSession aplicativo falhar ou o serviço Timeline estiver inativo, o formulário padrão será TezSession fechado após uma nova tentativa SessionPool
Correção de bugs	HIVE-23615: não desreferencia ponteiros nulos na classe de comandos do Beeline
Correção de bugs	HIVE-24849: criar um tempo limite de soquete de tabela externo quando o local tiver um grande número de arquivos (afeta 3.1.2)
Correção de bugs	HIVE-25209: consulta SELECT com função SUM produzindo resultado inesperado
Correção de bugs	HIVE-23666: A checkHashMode eficiência é ignorada quando um operador groupby não tem um conjunto de agrupamento
Correção de bugs	HIVE-23873: A consulta da tabela do Hive JDBCStorage Handler falha com o NPE quando o CBO está desativado

Tipo	Descrição
Correção de bugs	HIVE-24149: HiveStreamingConnection não fecha a conexão HMS
Correção de bugs	HIVE-25561: a tarefa encerrada não deve confirmar o arquivo. (afeta as versões 2.x e 3.x)
Correção de bugs	HIVE-25683: Feche o leitor na entrada. AcidUtils isRawFormatArquivo
Correção de bugs	HIVE-24294: as sessões podem gerar TezSessionPool AssertionError
Correção de bugs	HIVE-24182: Problema de autorização do Ranger com permanente UDFs
Correção de bugs	HIVE-22805: a vetorização com matriz ou mapa condicional não é implementada e gera um erro
Correção de bugs	HIVE-22828: Decimal64: instruções NVL e CASE convertem implicitamente decimal64 em 128
Correção de bugs	HIVE-21398: colunas com estatísticas estimadas não devem ser consideradas chaves exclusivas
Correção de bugs	HIVE-22490: adicionar arquivos jar com caracteres especiais no caminho gera um erro
Correção de bugs	HIVE-22700: as compactações podem vazar memória quando não autorizadas
Correção de bugs	HIVE-22053: o nome da função não é normalizado na criação da função

Tipo	Descrição
Correção de bugs	HIVE-22595: as inserções de partições dinâmicas apresentam falha na tabela Avro com esquema externo
Correção de bugs	HIVE-21795: a linha de resumo do rollup pode estar ausente quando um mapjoin está acontecendo em uma tabela particionada
Correção de bugs	HIVE-22987: ClassCastException em quando é nulo VectorCoalesce DataTypePhysicalVariation
Correção de bugs	HIVE-22219: desativar um gerenciador de nós bloqueia a reinicialização do serviço do LLAP
Correção de bugs	HIVE-21793: o CBO recuperará as estatísticas da coluna mesmo se hive.stats.fetch.column.stats estiver definido como false
Correção de bugs	HIVE-22163: CBO: habilitar o CBO ativa a estimativa de estatísticas, mesmo quando a estimativa está desabilitada
Correção de bugs	HIVE-18735: criar uma tabela similar perde o atributo transacional
Correção de bugs	HIVE-22433: Manipulador de armazenamento JDBC do Hive: resultados incorretos obtidos de BOOLEAN e TIMESTAMP da fonte de dados JDBC DataType
Correção de bugs	HIVE-19430:. ObjectStore cleanNotificationEvents OutOfMemory em um grande número de eventos pendentes
Correção de bugs	HIVE-20785: Nome de chave errado no JDBC. DatabaseMetaData getPrimaryKeys método

Tipo	Descrição
Correção de bugs	HIVE-16116: Beeline gera NPE quando beeline.hiveconfvariables={} em beeline.properties
Correção de bugs	HIVE-20066: hive.load.data.owner é comparado à entidade principal completa
Correção de bugs	HIVE-20489: explicar o plano de interrupções nas consultas
Correção de bugs	HIVE-21033: Esquecer de fechar a operação corta mais 2 saídas HiveServer
Correção de bugs	HIVE-19888: Aviso enganoso de “METASTORE_FILTER_HOOK será ignorado” de SessionState
Correção de bugs	HIVE-20303: INSERT OVERWRITE TABLE db.table PARTITION (...) IF NOT EXISTS lança InvalidTableException
Correção de bugs	HIVE-16144: CompactionInfo não tem equals/hashcode, mas é usado no Set
Correção de bugs	HIVE-20818: as vistas criadas com uma subconsulta WHERE considerarão as vistas referenciadas na subconsulta como entrada direta
Correção de bugs	HIVE-21005: LLAP: Lendo mais faixas por vazamento dividido ZlibCodecs
Correção de bugs	HIVE-20771: LazyBinarySerDe falha em estruturas vazias.
Correção de bugs	HIVE-18852: mensagem de erro enganosa na validação da tabela de alteração

Tipo	Descrição
Correção de bugs	HIVE-21124: HPL/SQL não é compatível com a instrução CREATE TABLE LIKE
Correção de bugs	HIVE-20935: O upload do tarball do pacote llap falha ao causar falha no início do serviço LLAP EC2
Correção de bugs	HIVE-20409: Hive ACID: Update/delete/merge não limpa o diretório de teste do hdfs
Correção de bugs	HIVE-20570: Union ALL com hive.optimize.union.remove=true tem um plano incorreto
Correção de bugs	HIVE-20421: entidade de caractere ilegal '\b' em hive-default.xml.template
Correção de bugs	HIVE-19133: As métricas de desempenho em fases do HS2 WebUI não são exibidas corretamente
Correção de bugs	HIVE-18977: listar partições retorna resultados diferentes com JDO e SQL direto
Correção de bugs	HIVE-20034: Reverta as alterações no tratamento de MetaStore exceções para fins de compatibilidade com versões anteriores
Correção de bugs	HIVE-20672: O encadeamento de login LlapTaskSchedulerService deve relatar cada intervalo fixo
Correção de bugs	HIVE-12812: habilitar mapred.input.dir.recursive por padrão para dar suporte à união com função agregada
Correção de bugs	HIVE-20147: a ingestão de streaming do Hive se satisaz com o registro sincronizado

Tipo	Descrição
Correção de bugs	HIVE-19203: Problema de segurança de rosca em HiveMetaStore
Correção de bugs	HIVE-20091: Tez: Adicione credenciais de segurança para saída FileSinkOperator
Correção de bugs	HIVE-16906: O Hive ATSHook deve verificar se há yarn.timeline-service.enabled antes de se conectar ao ATS
Correção de bugs	HIVE-20714: tblproperties SHOW para uma única propriedade retorna o valor na coluna de nome
Correção de bugs	HIVE-24730: as classes Shims substituem valores de hive-site.xml e tez-site.xml silencios amente
Correção de bugs	HIVE-22055: a contagem selecionada fornece um resultado incorreto após carregar dados do arquivo de texto

Amazon EMR 6.11.0 - Notas da versão do Hive

Amazon EMR 6.11.0 - Alterações no Hive

Tipo	Descrição
Melhoria	Adicionado suporte para descarte multithre aded de partições em vários segmentos para melhorar a peformance do descarte de partições
Melhoria	Compatível com a leitura de arquivos de consulta codificados do Hive

Tipo	Descrição
Melhoria	Habilitado o Tez Shuffle Handler por padrão para trabalhos do Hive no Tez
Bug	Adicionada uma opção para habilitar uma distribuição determinística de chaves para redutores na correção de resultados incorretos quando hive.groupby.skewindata está habilitado (informado no HIVE-20220)
Bug	Corrigida falha no cálculo de estatísticas quando o nome da partição padrão é configurado
Bug	Respeite todos os parâmetros de classificação SSL personalizados transmitidos quando o SSL é configurado imediatamente para HiveServer 2 em um cluster com criptografia em trânsito ativada
Backport	HIVE-23617: problemas de API de armazenamento corrigidos FindBug
Backport	HIVE-26408: Vetorização: corrija a desalocação de colunas de rascunho, não reutilize uma criança como saída ConstantVectorExpression
Backport	HIVE-23614: Sempre passe para HiveConfig removeTempOr DuplicateFiles
Backport	HIVE-23354: Remova a verificação de integridade do tamanho do arquivo de compareTempOr DuplicateFiles

Tipo	Descrição
Backport	HIVE-20344: corrigido PrivilegeSynchroni zer para lançamento de SBA. AccessCon trolException Também introduzida a proprieda de hive.privilege.synchronizer para desabilitar o sincronizador de privilégios
Backport	HIVE-15826: Support a configuração de 'serialization.encoding' para todos SerDes
Backport	HIVE-18284: corrigir o NPE ao inserir dados com a cláusula ‘distribuir por’ com otimização de classificação dynpart
Backport	HIVE-24930: o curto-circuito Operator. setDone() da operação filha não é usado no caminho de código vetorizado (se childSize == 1)
Backport	HIVE-24523: O caminho de leitura vetorizad o para LazySimpleSerde não respeita o SERDEPROPERTIES para timestamp
Backport	HIVE-23265: conjuntos de linhas duplicadas são retornados com os conjuntos Limit e Offset
Backport	HIVE-21492: não VectorizedParquetR ecordReader consigo ler o arquivo de parquete gerado usando uma ferramenta econômica ou personalizada
Backport	HIVE-22540: Vetorização: colunas decimal64 não funcionam com. VectorizedBatchUtil makeLikeColumnVetor ()
Backport	HIVE-22588: descarregar as linhas restantes para o resto dos conjuntos de agrupamento ao mudar o modo groupby vetorial

Tipo	Descrição
Backport	HIVE-22551: O BytesColumnVector initBuffer deve limpar o vetor e o comprimento de forma consistente
Backport	HIVE-22448: CBO: expanda a contagem múltipla distinta com uma chave de agrupamento
Backport	HIVE-22248: corrigir problemas persistentes das estatísticas
Backport	HIVE-22210: a vetorização pode reutilizar colunas de saída de computação envolvidas na filtragem
Backport	HIVE-21531: vetorização: todos os hashcodes NULL não são computados usando Murmur3
Backport	HIVE-20419: Vetorização: evite a mutação após ser usado em uma chave de hashmap VectorPartitionDesc
Backport	HIVE-19388: durante a inicialização ClassCast Exception VectorMapJoinCommonOperator
Backport	HIVE-21584: Preparação para o Java 11: o carregador de classes do sistema não é o carregador URLClass
Backport	HIVE-25107: o registro em log do caminho de classe deve estar no nível DEBUG (#2271)
Backport	HIVE-22097: java.util incompatível. ArrayList para java 11

Tipo	Descrição
Backport	HIVE-23938: LLAP: JDK11 - alguns argumentos jvm relacionados à rotação do arquivo de log do GC não podem mais ser usados
Backport	HIVE-26226: excluir jdk.tools dep do hive-metastore em upgrade-acid
Backport	HIVE-17879: atualizar o plug-in Datanucleus Maven
Backport	HIVE-27004: DateTimeFormatterBuilder # appendZoneText não pode analisar 'UTC+' em versões Java superiores a 8
Backport	HIVE-16812: VectorizedOrcAcidRowBatchReader não filtra eventos de exclusão
Backport	HIVE-17917:. VectorizedOrcAcidRowBatchReader computeOffsetAndOtimização do bucket
Backport	HIVE-19985: ACID: ignorar a decodificação das seções ROW__ID para consultas somente leitura
Backport	HIVE-20635: VectorizedOrcAcidRowBatchReader não filtra eventos de exclusão dos arquivos originais
Upgrade	Atualizar o Javadoc para 3.3.1
Upgrade	Atualizar o Javassist para 3.24.1-GA
Upgrade	Atualização apache-directory-server para 2.0.0-M14

Novas configurações

Name	Classificação	Descrição
hive.metastore.fs.drop.partition.threads	hive-site	Número de threads principais no grupo de threads de partições descartadas.
hive.metastore.fs.drop.partition.keepalive.time	hive-site	Tempo em segundos em que um thread assíncrono de partições descartadas ociosas (do grupo de threads) aguardará a chegada de uma nova tarefa antes de ser encerrado.
hive.metastore.fs.drop.partition.threadpool.max.queue.size	hive-site	Tamanho máximo da fila a ser usada no grupo de threads para descartar partições do sistema de arquivos.
hive.groupby.enable.deterministic.distribution	hive-site	Habilite a distribuição determinística de chaves para redutores. Isso passará um valor de semente constante ao chamar a função rand usada para particionamento aleatório.
hive.privilege.synchronizer	hive-site	Se os privilégios do autorizador externo devem ser sincronizados periodicamente em 2. HiveServer
hive.cli.query.file.encoding	hive-site	Codificação de arquivo para todos os tipos de arquivos de consulta (arquivo de consulta, arquivo de consulta inicial,

Name	Classificação	Descrição
		arquivo rc etc.) fornecidos nos argumentos da cli.
hive.emr.tez.shuffle.enabled	hive-site	Os trabalhos do Hive no Tez agora usam tez_shuffle por padrão em vez de mapreduce _shuffle como Shuffle Handler padrão.

Configurações descontinuadas

As propriedades de configuração a seguir foram descontinuadas como resultado do HIVE-23354 e não são mais compatíveis com as versões 6.11.0 e posteriores do Amazon EMR.

Name	Valor padrão
`hive.mapred.reduce.tasks.speculative.execution`	`false`
`tez.am.speculation.enabled`	`false`

Amazon EMR 6.10.0 - Notas da versão do Hive

Amazon EMR 6.10.0 - Alterações no Hive

Tipo	Descrição
Recurso	Ative controles de acesso AWS Lake Formation baseados para consultas do Apache Hive (gravação) por meio do IAM Passthrough (API HiveCli/Steps).
Melhoria	Desativar a configuração `hive.log.explain.output` por padrão para reduzir o tamanho do log

Tipo	Descrição
Backport	HIVE-26408: Vetorização: corrija a desalocação de colunas de rascunho, não reutilize uma criança como saída ConstantVectorExpression
Backport	HIVE-22269: corrigir a contagem incorreta de redutores nas consultas de inserção com partição dinâmica devido à falta de estatísticas causada pelo HIVE-20703.
Backport	HIVE-22891: Ignore a PartitionDesc extração para o modo de execução não LLAP CombineHiveRecord
Backport	HIVE-23804: adicionar banco de dados padrão para tabelas específicas de estatísticas de colunas no esquema de metastore do Hive para torná-las compatíveis retroativamente
Backport	HIVE-25277: Exclusão lenta de partições do Hive para armazenamentos de objetos em nuvem com custos caros ListFiles
Backport	HIVE-19202: O CBO falhou devido à entrada. NullPointerException HiveAggregate isBucketedInput()
Backport	HIVE-19048: corrigir erros Initscript do Beeline são ignorados
Backport	HIVE-21085: registro de visões materializadas inicia sessão do tez não externa
Backport	HIVE-21675: CREATE VIEW IF NOT EXISTS retorna um erro em vez de “OK” se a visualização já existir. Esta é uma regressão do Hive 2.

Tipo	Descrição
Backport	HIVE-21646: Sim: Evite TezTasks escapar do contexto de registro de threads
Backport	HIVE-22054: evitar listagem recursiva para verificar se um diretório está vazio
Backport	HIVE-16587: NPE ao inserir tipos complexos com valores nulos aninhados
Backport	HIVE-22647: habilitar o conjunto de sessões por padrão
Backport	HIVE-13288: Mensagem de exceção confusa em .localizeResource DagUtils
Backport	HIVE-23870: Otimize várias conversões de texto em. WritableHiveCharObjectInspector getPrimitiveJavaObjeto/ HiveCharWritable
Backport	HIVE-21498: atualizar o Thrift para 0.13.0
Backport	HIVE-24378: espaços iniciais e finais não são removidos antes da conversão decimal
Backport	HIVE-21341: padrões sensatos: hive.server2.idle.operation.timeout e hive.server2.idle.session.timeout são muito altos
Backport	HIVE-22465: Adicionar configuração ssl em TezConfigurationFactory
Backport	HIVE-24710: otimizar a iteração de PTF para contagem (*) para reduzir o custo de CPU e E/S
Backport	HIVE-15406: considere a vetorização da nova função 'trunc'

Tipo	Descrição
Backport	HIVE-21541: corrigir cabeçalhos asf ausentes do HIVE-15406
Backport	HIVE-24808: datas analisadas em cache
Backport	HIVE-24746: PTF: TimestampValueBoundaryScanner pode ser otimizado durante o cálculo do alcance
Backport	HIVE-25059: o evento Alter é convertido em renomeação durante a replicação
Backport	HIVE-25142: refazer o hash na tabela de hash rápida do mapa, causando corrupção em chaves grandes
Backport	HIVE-23756: adicionadas mais restrições ao arquivo package.jdo
Backport	HIVE-25150: os caracteres de tabulação não são removidos antes da conversão decimal de maneira similar ao caractere de espaço que é corrigido como parte do HIVE-24378
Backport	HIVE-25093: date_format() UDF está retornando a saída somente no fuso horário UTC
Backport	HIVE-25268: date_format udf retorna resultados incorretos para datas anteriores a 1900 se o fuso horário local é diferente de UTC
Backport	HIVE-25338: AIOBE em UDF de configuração se a entrada estiver vazia
Backport	HIVE-22400: o minuto da UDF com o tempo retorna NULL

Tipo	Descrição
Backport	HIVE-25058: PTF: TimestampValueBoundaryScanner pode ser otimizado durante o cálculo do alcance pt2 - isDistanceGreater
Backport	HIVE-25449: datediff() fornece uma saída errada quando executado em uma tarefa do tez com algum fuso horário não UTC
Backport	HIVE-23688: Vetorização: IndexArrayOutOfBoundsException para coluna de tipo de mapa que inclui valor nulo
Backport	HIVE-22247: O Hive é lançado quando a saída da tarefa da HFile OutputFormat partição está FileNotFoundException vazia
Backport	HIVE-25570: o Hive deve enviar o caminho de URL completo para autorização do comando insert overwrite location
Backport	HIVE-22903: row_number() vetorizado redefine o número da linha após um lote no caso de expressão constante em cláusula de partição
Backport	HIVE-25549: resultados incorretos para a função de janela com expressão na cláusula PARTITION BY ou ORDER BY
Backport	HIVE-25579: a substituição de LOAD acrescenta em vez de substituir
Backport	HIVE-25659: consultas sql diretas do metastore com IN/(NOT IN) devem ser divididas com base nos parâmetros máximos permitidos por SQL DB

Tipo	Descrição
Backport	HIVE-20502: corrigir o NPE ao executar skewjoin_mapjoin10.q quando as estatísticas da coluna são usadas.
Backport	HIVE-25765: a propriedade skip.header.line.count ignora as linhas de cada bloco quando o tamanho do arquivo é maior FetchOperator
Bug	Corrigir o NPE na inserção em determinados cenários quando `hive.stats.column.autogather` e `hive.groupby.skewindata` estão habilitados
Bug	Corrigir o NPE quando o valor `mapred.tasktracker.expiry.interval` não está definido

Amazon EMR 6.9.0 - Notas da versão do Hive

Amazon EMR 6.9.0 - Alterações no Hive

Tipo	Descrição
Upgrade	Atualizar o Jetty para 9.4.48.v20220622
Upgrade	Suporte para Hadoop 3.3.3
Recurso	Integração do Hive no Amazon EMR com o Lake Formation para workloads interativas usando a API do GCSC.
Recurso	Integração do Hive no Amazon EMR com o Iceberg.
Melhoria	Ative o SSL em HiveServer 2 quando a criptografia em trânsito estiver habilitada

Tipo	Descrição
	usando as configurações de segurança do Amazon EMR.
Melhoria	Habilite o confirmador otimizado para EMRFS Amazon S3 do Hive por padrão. Para obter mais informações, consulte Habilitar o confirmador otimizado para EMRFS S3 do Hive.
Melhoria	Adicione `HiveHBaseTableInputFormatV2` que herda apenas a versão mapeada do InputFormat para corrigir o SPARK-34210. Defina `hive.hbase.inputformat.v2` como `true` para usá-lo.
Melhoria	Espere que o TeZam seja iniciado em segundo plano com hive.cli.tez.session.async em vez de encerrá-lo e iniciar um novo imediatamente. Use `hive.emr.cli.tez.session.open.timeout` para definir esse tempo limite em segundos.
Melhoria	Adicione a opção hive.conf.restricted.list.append para anexar configurações separadas por vírgula à lista de configurações restritas existente. `hive.conf.restricted.list`
Melhoria	Mensagem de erro mais clara quando a consulta do Hive apresenta falha porque o local não está definido para o banco de dados.
Backport	HIVE-24484: atualizar o Hadoop para 3.3.1 e o Tez para 0.10.2
Backport	HIVE-22398: Remova o gerenciamento de filas do YARN via. ShimLoader

Tipo	Descrição
Backport	HIVE-23190: LLAP: modifique IndexCache para passar o objeto do sistema de arquivos para o. TezSpillRecord
Backport	HIVE-22185: O HADOOP-15832 causará problemas com testes usando clusters. MiniYarn
Backport	HIVE-21670: substituir `mockito-all` pela dependência `mockito-core` .
Backport	HIVE-24542: preparar o Guava para atualizaç ões.
Backport	HIVE-23751: QTest: Substitua o `#mkdirs()` método para alinhar após o HADOOP-16582. `ProxyFileSystem`
Backport	HIVE-21603: preparação do Java 11: atualizar a versão powermock.
Backport	HIVE-24083: erro `hcatalog` no Hadoop 3.3.0: tipo de autenticação necessário.
Backport	HIVE-24282: mostrar colunas não deve classificar as colunas de saída, a menos que seja explicitamente mencionado.
Backport	HIVE-20656: padrões sensatos: as configura ções de memória de agregação de mapas são muito agressivas.
Backport	HIVE-25443: A seta SerDe não pode serializar/ desserializar tipos de dados complexos quando há mais de 1024 valores

Tipo	Descrição
Backport	HIVE-19792: atualizar o orc para 1.5.2 e habilitar os testes de evolução do esquema decimal_64.
Backport	HIVE-20437: manipular a evolução do esquema com base em flutuação, duplo e decimal.
Backport	HIVE-21987: o Hive não consegue ler o Parquet int32 anotado com decimal.
Backport	HIVE-20038: atualizar consultas em tabelas sem buckets e particionadas gera NPE.

Amazon EMR 6.9.0 - Problemas conhecidos do Hive

- Com o Amazon EMR 6.6.0 a 6.9.x, as consultas INSERT com partição dinâmica e uma cláusula ORDER BY ou SORT BY sempre terá dois redutores. Esse problema é causado pela alteração do OSS HIVE-20703, que coloca a otimização da partição dinâmica de classificação sob uma decisão baseada em custos. Se sua workload não exigir a classificação de partições dinâmicas, recomendamos que você defina a propriedade `hive.optimize.sort.dynamic.partition.threshold` como `-1` para desabilitar o novo atributo e obter o número de redutores calculado corretamente. Esse problema foi corrigido no OSS Hive como parte do HIVE-22269 e foi corrigido no Amazon EMR 6.10.0.

Amazon EMR 6.8.0 - Notas da versão do Hive

Amazon EMR 6.8.0 - Alterações no Hive

Tipo	Descrição
Melhoria	Reduza as chamadas do sistema de arquivos no comando msck. Melhorias de performance (~ 15 a 20 x em mais de 10 mil partições)

Tipo	Descrição
Backport	HIVE-20678: O Hive HBase TableOutp utFormat deve ser implementado para garantir a compatibilidade HiveOutputFormat
Backport	HIVE-21040: o msck faz uma listagem desnecessária de arquivos no último nível da árvore de diretórios
Backport	HIVE-21460: carregar dados seguidos por uma consulta * selecionada resulta em resultados incorretos
Backport	HIVE-21660: resultado incorreto quando union all e a vista posterior com explosão são usadas
Backport	HIVE-22505: ClassCastException causado pela seleção incorreta do operador vetorizado
Backport	HIVE-22513: a propagação constante da coluna convertida em operações de filtro pode causar resultados incorretos
Backport	HIVE-23435: o resultado completo da junção externa é composto de linhas ausentes
Backport	HIVE-24209: conversão incorreta do argumento de pesquisa para a operação NOT BETWEEN quando a vetorização está habilitada
Backport	HIVE-24934: a VectorizedExpressions anotação não é necessária na verificação genérica UDFSQCount
Backport	HIVE-25278: HiveProjectJoinTransposeRul e pode fazer transformações inválidas com expressões de janela

Tipo	Descrição
Backport	HIVE-25505: resultados incorretos com cabeçalho. skip.header.line.count se a primeira linha está em branco
Backport	HIVE-26080: atualizar o accumulo-core para 1.10.1
Backport	HIVE-26235: a condição OR na coluna binária está retornando um resultado vazio
Bug	Corrige vários registros de aviso de ligações SLF4 J no stderr durante o lançamento
Bug	Corrija a falha na consulta SHOW TABLE EXTENDED com erro FS errado quando a partição e a tabela estão em sistemas de arquivos diferentes.

Amazon EMR 6.8.0 - Problemas conhecidos do Hive

- Com o Amazon EMR 6.6.0 a 6.9.x, as consultas INSERT com partição dinâmica e uma cláusula ORDER BY ou SORT BY sempre terá dois redutores. Esse problema é causado pela alteração do OSS HIVE-20703, que coloca a otimização da partição dinâmica de classificação sob uma decisão baseada em custos. Se sua workload não exigir a classificação de partições dinâmicas, recomendamos que você defina a propriedade `hive.optimize.sort.dynamic.partition.threshold` como `-1` para desabilitar o novo atributo e obter o número de redutores calculado corretamente. Esse problema foi corrigido no OSS Hive como parte do HIVE-22269 e foi corrigido no Amazon EMR 6.10.0.

Amazon EMR 6.7.0 - Notas da versão do Hive

Amazon EMR 6.7.0 - Alterações no Hive

Tipo	Descrição
Recurso	Integração do Amazon EMR Hive com. LakeFormation
Recurso	Registro em log de auditoria adicional para o confirmador otimizado para EMRFS Amazon S3 do Hive. Configuração do Hive: `hive.blobstore.output-committer.logging` , padrão: `false`
Recurso	Excluído o diretório de destino na instrução insert overdrive com um resultado de seleção vazio em uma tabela/partição estática não particionada para se comportar de forma semelhante ao Hive 2.x. Configuração do Hive: `hive.emr.iow.clean.target.dir` , padrão: `false`
Bug	Corrigida falha de consulta intermitente ao ser usado o confirmador otimizado para EMRFS Amazon S3 do Hive com classificação de buckets de partição.
Upgrade	Versão do Hive atualizada para 3.1.3. Consulte as notas de lançamento do Apache Hive 3.1.3 para obter mais detalhes.
Upgrade	Atualizado o Parquet para 1.12.2.
Backport	HIVE-20065: o metastore não deve contar com o jackson 1.x
Backport	HIVE-20071: migrar para o jackson 2.x e evitar o uso

Tipo	Descrição
Backport	HIVE-20607: TxnHandler deve ser usado PreparedStatement para executar consultas SQL diretas
Backport	HIVE-20740: Remova o bloqueio global no método.setConf ObjectStore
Backport	HIVE-20961: retirar a implementação do NVL
Backport	HIVE-22059: o jar hive-exec não contém a biblioteca jackson (fasterxml)
Backport	HIVE-22351: Corrige o uso incorreto do encadeamento em ObjectStore TestObjec tStore
Backport	HIVE-23534: NPE em RetryingMetaStoreC lient #invoke ao capturar sem mensagem MetaException
Backport	HIVE-24048: harmonizar os componentes do Jackson com a versão 2.10.latest - Hive
Backport	HIVE-24768: usar jackson-bom em todos os lugares para a substituição da versão
Backport	HIVE-24816: Atualize jackson para 2.10.5.1 ou 2.11.0+ devido ao CVE-2020-25649
Backport	HIVE-25971: o desligamento da tarefa Tez sendo atrasado devido ao pool de threads em cache não fechado
Backport	HIVE-26036: NPE causado por get () em MTable ObjectStore

Amazon EMR 6.7.0 - Problemas conhecidos do Hive

- Consultas com funções de janela na mesma coluna da junção podem levar a transformações inválidas, conforme informado no HIVE-25278, e causar resultados incorretos ou falhas nas consultas. Uma solução alternativa seria desabilitar o CBO em nível de consulta para essas consultas. A correção estará disponível em uma versão do Amazon EMR após a 6.7.0. Para obter mais informações, entre em contato com AWS o suporte.
- Com o Amazon EMR 6.6.0 a 6.9.x, as consultas INSERT com partição dinâmica e uma cláusula ORDER BY ou SORT BY sempre terá dois redutores. Esse problema é causado pela alteração do OSS HIVE-20703, que coloca a otimização da partição dinâmica de classificação sob uma decisão baseada em custos. Se sua workload não exigir a classificação de partições dinâmicas, recomendamos que você defina a propriedade `hive.optimize.sort.dynamic.partition.threshold` como `-1` para desabilitar o novo atributo e obter o número de redutores calculado corretamente. Esse problema foi corrigido no OSS Hive como parte do HIVE-22269 e foi corrigido no Amazon EMR 6.10.0.

Amazon EMR 6.6.0 - Notas da versão do Hive

Amazon EMR 6.6.0 - Alterações no Hive

Tipo	Descrição
Upgrade	Atualize o Parquet para 1.12.1.
Upgrade	Atualizar a versão de jetty jars para 9.4.43.v2 0210629
Bug	Corrigido um problema que fazia com que o Hive fosse instalado em todos os nós centrais/ de tarefa quando LLAP estava habilitado em um cluster do Hive.
Backport	HIVE-25942: atualizar o commons-io para 2.8.0 devido ao CVE-2021-29425
Backport	HIVE-25726: atualizar a velocidade para 2.3 devido ao CVE-2020-13936

Tipo	Descrição
Backport	HIVE-25680: Autorize a API HiveMetastore do servidor #get_table_meta a usar qualquer modelo de autorização. HiveMetastore
Backport	HIVE-25554: atualizar a versão do Arrow para 0.15
Backport	HIVE-25242: a consulta é executada de forma extremamente lenta com vectorized.adaptor = chosen
Backport	HIVE-25085: MetaStore Os clientes não são mais compartilhados entre as sessões.
Backport	HIVE-24827: a consulta de agregação do Hive retorna resultados incorretos para arquivos que não sejam de texto.
Backport	HIVE-24683: getFileId Hadoop23Shims propenso a NPE para caminhos inexistentes
Backport	HIVE-24656: o CBO falha em consultas com is null em tipos de mapa e matriz
Backport	HIVE-24556: Otimize DefaultGraphWalker para casos sem neto
Backport	HIVE-24408: atualizar o Parquet para 1.11.1
Backport	HIVE-24391: Corrigir falhas de correção no branch-3.1 TestOrcFile
Backport	HIVE-24362: o processamento da árvore AST é abaixo do ideal para árvores com grande número de nós

Tipo	Descrição
Backport	HIVE-24316: atualizar o ORC de 1.5.6 para 1.5.8 em branch-3.1
Backport	HIVE-24307: Beeline com property-file e -e parameter está apresentando falha
Backport	HIVE-24245: PTF vetorizado com contagem e over partition distinta produzindo resultados incorretos.
Backport	HIVE-24224: corrigir a ação de ignorar o cabeçalho/rodapé do Hive no Tez em arquivo compactado
Backport	HIVE-24157: modo estrito de falha em CAST timestamp ↔ numérico
Backport	HIVE-24113: NPE em genérico UDFTo UnixTimeStamp
Backport	HIVE-23987: atualizar a versão do Arrow para 0.11.0
Backport	HIVE-23972: adicionar ID de cliente externo ao cliente externo do LLAP
Backport	HIVE-23806: evitar limpar os estados de estatísticas da coluna em todas as partições caso o esquema seja estendido. Isso melhora o runtime da instrução alter table add columns.
Backport	HIVE-23779: As BasicStatsTask informações não estão sendo impressas no console beeline
Backport	HIVE-23306: o comando RESET não funciona se há uma configuração definida por System.getProperty

Tipo	Descrição
Backport	HIVE-23164: o servidor não foi encerrado corretamente devido a threads que não são do daemon
Backport	HIVE-22967: suporte para hive.reloadable.aux.jars.path para Hive no Tez
Backport	HIVE-22934: contadores de log interativos do servidor do Hive para fluxo de erros
Backport	HIVE-22901: a substituição de variáveis pode levar à OOM em referências circulares
Backport	HIVE-22769: resultados incorretos de consultas e falha nas consultas durante a geração dividida para arquivos de texto compactados
Backport	HIVE-22716: A leitura para está interrompida ByteBuffer ParquetFooterInputFromCache
Backport	HIVE-22648: atualizar o Parquet para 1.11.0
Backport	HIVE-22640: Decimal64ColumnVector: ClassCastException quando o tipo de coluna de partição é Decimal
Backport	HIVE-22621: caso de teste instável: .testSigning TestLlapSignerImpl
Backport	HIVE-22533: corrigir possíveis vulnerabilidades da interface de usuário da Web do daemon do LLAP
Backport	HIVE-22532: O PTFPPD pode ultrapassar o limite incorretamente por meio da classificação/função DenseRank

Tipo	Descrição
Backport	HIVE-22514: HiveProtoLoggingHook pode consumir muita memória
Backport	HIVE-22476: a função datediff do Hive forneceu resultados inconsistentes quando hive.fetch.task.conversion foi definida como nenhuma
Backport	HIVE-22429: tabelas em cluster migradas usando bucketing_version 1 no hive 3 usam bucketing_version 2 para inserções
Backport	HIVE-22412: StatsUtils lance NPE ao explicar
Backport	HIVE-22360: MultiDelimitSerDe retorna resultados errados na última coluna quando o arquivo carregado tem mais colunas do que aquelas no esquema da tabela
Backport	HIVE-22332: o Hive deve garantir configurações válidas de evolução do esquema desde o ORC-540
Backport	HIVE-22331: unix_timestamp sem argumento retorna carimbo de data/hora em milissegundos em vez de em segundos
Backport	HIVE-22275:. OperationManager queryIdOperationnão limpa adequadamente vários QueryIDs
Backport	HIVE-22273: a verificação de acesso apresenta falha quando um diretório temporário é removido
Backport	HIVE-22270: atualizar o commons-io para 2.6

Tipo	Descrição
Backport	HIVE-22241: implementar UDF para interpret ar a data/carimbo de data/hora usando sua representação interna e o calendário híbrido gregoriano-juliano
Backport	HIVE-22241: implementar UDF para interpret ar a data/carimbo de data/hora usando sua representação interna e gregoriano-juliano híbrido
Backport	HIVE-22232: NPE quando hive.order.columna lignment é definido como falso
Backport	HIVE-22231: a consulta do Hive com tamanho grande por meio do knox apresenta falha na gravação de pipe defeituoso
Backport	HIVE-22221: Cliente externo Llap - É necessári o reduzir #getSplits LlapBaseInputFormat
Backport	HIVE-22208: o nome da coluna com a palavra-chave reservada não escapa quando consultas que incluem junção na tabela com a coluna de máscara é regravada
Backport	HIVE-22197: junção de mesclagem comum gerando exceção de conversão de classe.
Backport	HIVE-22170: from_unixtime e unix_timestamp devem usar o fuso horário da sessão do usuário
Backport	HIVE-22169: Tez: SplitGenerator tenta procurar arquivos de plano que não existirão para Tez
Backport	HIVE-22168: remover registros em log muito caros do hotpath de cache do llap

Tipo	Descrição
Backport	HIVE-22161: UDF: sincroniza em org.apach e.hadoop.hive.ql.udf. FunctionRegistry UDFType classe
Backport	HIVE-22120: Corrige ArrayOutOfBound resultados/exceções errados nas junções do mapa externo esquerdo em condições de limite específicas
Backport	HIVE-22115: impedir a criação de anexador de roteamento de consultas se a propriedade é definida como falsa
Backport	HIVE-22113: Evitar o desligamento do LLAP em casos relacionados AMReporter RuntimeEx ception
Backport	HIVE-22106: remover a sincronização de consultas cruzadas para a avaliação de partição
Backport	HIVE-22099: Várias datas relacionadas não UDFs conseguem lidar com datas julianas adequadamente desde HIVE-20007
Backport	HIVE-22037: HS2 deve registrar ao desligar devido ao OOM
Backport	HIVE-21976: O offset deve ser nulo em vez de zero em Calcite HiveSortLimit
Backport	HIVE-21924: dividir arquivos de texto mesmo que exista cabeçalho/rodapé
Backport	HIVE-21913: UDTFGet Divisões genéricas devem lidar com nomes de usuário da mesma forma que o LLAP

Tipo	Descrição
Backport	HIVE-21905: Melhoria dos genéricos em toda a classe FetchOperator
Backport	HIVE-21902: HiveServer 2 UI: necessidades do cabeçalho de resposta do cais X-Frame-O ptions
Backport	HIVE-21888: definir o padrão hive.parq uet.timestamp.skip.conversion como verdadeiro
Backport	HIVE-21868: vetorizar CAST...FORMAT
Backport	HIVE-21864: #closeAll LlapBaseInputFormat
Backport	HIVE-21863: melhorar a conversão do tipo vetorizador para a expressão WHEN
Backport	HIVE-21862: ORC ppd produz resultado errado com carimbo de data/hora
Backport	HIVE-21846: Crie um tópico no TeZam que busque métricas periodicamente LlapDaemon
Backport	HIVE-21837: MapJoin está lançando uma exceção quando a coluna selecionada tem valores completamente nulos
Backport	HIVE-21834: evitar chamadas desnecessárias para simplificar as condições do filtro
Backport	HIVE-21832: Novas métricas para obter o tempo médio queue/serving/response
Backport	HIVE-21827: Várias chamadas no método SemanticAnalyzer não passam pelo getTableO bject ByName

Tipo	Descrição
Backport	HIVE-21822: Exponha LlapDaemon métricas por meio de um novo método de API
Backport	HIVE-21818: CBO: A cópia tem tráfego de metastore TableRelOptHiveTable
Backport	HIVE-21815: as estatísticas no arquivo ORC são analisadas duas vezes
Backport	HIVE-21805: HiveServer 2: Use o rápido ShutdownHookManager APIs
Backport	HIVE-21799: NullPointerException em DynamicPartitionPruningOptimization, quando a chave de junção está na coluna de agregação
Backport	HIVE-21794: Adicione parâmetros de visualização materializada ao sqlStdAuth SafeVarNameRegexes
Backport	HIVE-21768: JDBC: remover o prefixo de união padrão para consultas UNION não incluídas
Backport	HIVE-21746: ArrayIndexOutOfBoundsException durante a junção de hash particionada dinamicamente, com o CBO desativado
Backport	HIVE-21717: a renomeação do diretório está falhando na tarefa de movimentação.
Backport	HIVE-21685: simplificação incorreta na consulta com várias cláusulas IN
Backport	HIVE-21681: descrever formatado mostra informações incorretas para várias chaves primárias

Tipo	Descrição
Backport	HIVE-21651: mover serde de protobuf (buffer de protocolo) para hive-exec.
Backport	HIVE-21619: explicação estendida sobre imprimir o tipo de carimbo de data/hora sem precisão no SQL
Backport	HIVE-21592: não OptimizedSql é mostrado quando a expressão contém CONCAT
Backport	HIVE-21576: apresentar CAST...FORMAT e uma lista limitada de formatos de data e hora do SQL:2016
Backport	HIVE-21573: o transporte binário deve ignorar a entidade principal se a autenticação está definida como delegationToken
Backport	HIVE-21550: TestObjectStore os testes são instáveis - Não foi possível obter um bloqueio dentro do tempo solicitado
Backport	HIVE-21544: A propagação constante corrompe as expressões durante o dobramento coalesce/case/when
Backport	HIVE-21539: GroupBy + onde a cláusula na mesma coluna resulta em reescrita incorreta da consulta
Backport	HIVE-21538: Beeline: fonte da senha, embora o leitor do console não passe para o parâmetro de conexão
Backport	HIVE-21509: o LLAP pode armazenar em cache vetores de coluna corrompidos e retornar um resultado de consulta incorreto

Tipo	Descrição
Backport	HIVE-21499: não deve remover a função do registro se o comando de criação falhar com AlreadyExistsException
Backport	HIVE-21496: o dimensionamento automático de buffer não ordenado pode estourar
Backport	HIVE-21468: diferenciação entre maiúsculas e minúsculas em nomes de identificadores para o manipulador de armazenamento JDBC
Backport	HIVE-21467: remover as importações junit.framework.Assert desativadas
Backport	HIVE-21435: LlapBaseInputFormat deve obter o número da tarefa da configuração TASK_ATTEMPT_ID, se presente, durante a construção SubmitWorkRequestProto
Backport	HIVE-21389: a distribuição do Hive perde javax.ws.rs-api.jar após o HIVE-21247
Backport	HIVE-21385: permitir a desativação do pushdown de computação não divisível para fontes JDBC
Backport	HIVE-21383: manipulador de armazenamento JDBC: usar catálogo e esquema para recuperar tabelas, se especificado
Backport	HIVE-21382: otimização de redução de grupos por chaves - as chaves não são reduzidas em query23
Backport	HIVE-21362: adicionar um formato de entrada e serde para ler arquivos protobuf.

Tipo	Descrição
Backport	HIVE-21340: CBO: podar colunas não chave que alimentam um SemiJoin
Backport	HIVE-21332: limpar os buffers não bloqueados em vez dos bloqueados
Backport	HIVE-21329: tamanho do buffer de saída não ordenado no runtime do Tez personalizado, dependendo do pipeline do operador
Backport	HIVE-21295: StorageHandler deve converter a data em string usando a convenção Hive
Backport	HIVE-21294: vetorização: o 1-reducer Shuffle pode ignorar as funções de hash de objeto
Backport	HIVE-21255: Remover em QueryConditionBuil der JdbcStorageHandler
Backport	HIVE-21253: Support em JDBC DB2 StorageHandler
Backport	HIVE-21232: LLAP: adicionar um provedor de afinidade dividida compatível com perda de cache
Backport	HIVE-21214: MoveTask : Use AttemptID em vez do tamanho do arquivo para desduplicação de arquivos compareTempOr DuplicateFiles
Backport	HIVE-21184: adicionar explicação e explicar o plano CBO formatado com informações de custo
Backport	HIVE-21182: ignorar a configuração do diretório temporário do hive durante o planejamento

Tipo	Descrição
Backport	HIVE-21171: ignorar a criação de diretórios temporários para o tez se o RPC está ativado
Backport	HIVE-21126: Permitir consultas em nível de sessão em #getSplit LlapBaseInputFormat
Backport	HIVE-21107: erro “Cannot find field” durante a junção de hash particionada dinamicamente
Backport	HIVE-21061: A consulta CTAS falha com uma fonte vazia IllegalStateException
Backport	HIVE-21041: NPE, ParseException na obtenção do esquema do plano lógico
Backport	HIVE-21013: JdbcStorageHandler falha ao encontrar a coluna de partição no Oracle
Backport	HIVE-21006: Estenda SharedWorkOptimizer para remover semijunções quando houver uma oportunidade de reutilização
Backport	HIVE-20992: dividir a configuração hive.metastore.dbaccess.ssl.properties em configurações mais significativas
Backport	HIVE-20989: JDBC - O log GetOperationStatus + pode bloquear o progresso da consulta via hibernação
Backport	HIVE-20988: resultados incorretos para consultas agrupadas com chave primária em várias colunas

Tipo	Descrição
Backport	HIVE-20985: se as entradas do operador selecionado são colunas temporárias, a vetorização pode reutilizar algumas delas como saída
Backport	HIVE-20978: “hive.jdbc.*” deve ser adicionado ao sqlStdAuth SafeVarNameRegexes
Backport	HIVE-20953: remover uma função do registro de funções quando ela não pode ser adicionad a ao metastore ao ser criada.
Backport	HIVE-20952: Limpando .java Vectoriza tionContext
Backport	HIVE-20951: LLAP: definir Xms como 50% sempre
Backport	HIVE-20949: melhorar a estimativa da cardinali dade de PKFK no planejamento físico
Backport	HIVE-20944: não validar as estatísticas durante a compilação da consulta
Backport	HIVE-20940: casos de transição em que a resolução de tipo do Calcite é mais estrita do que o Hive.
Backport	HIVE-20937: falha na consulta jdbc do Postgres com “LIMIT must not be negative”
Backport	HIVE-20926: a dica de redução de semijunção apresenta falha quando as entradas do filtro de bloom são elevadas ou quando não há estatísti cas

Tipo	Descrição
Backport	HIVE-20920: usar restrições SQL para melhorar o algoritmo de reordenação de junções
Backport	HIVE-20918: sinalizador para habilitar/desabilitar o pushdown de computação do Calcite para a conexão JDBC
Backport	HIVE-20915: disponibilizar a otimização dinâmica da partição de classificação para HoS e MR
Backport	HIVE-20910: a inserção em uma tabela com buckets apresenta falha devido à otimização dinâmica da classificação de partições
Backport	HIVE-20899: o URI do Keytab para o serviço YARN do LLAP é restritivo para oferecer suporte somente a HDFS
Backport	HIVE-20898: para funções relacionadas ao tempo, os argumentos não podem ser convertidos em um tipo não anulável
Backport	HIVE-20881: a propagação constante simplifica demais as projeções
Backport	HIVE-20880: atualizar o valor padrão de hive.stats.filter.in.min.ratio
Backport	HIVE-20873: Use o hash Murmur para reduzir a colisão de hash VectorHashKeyWrapperTwoLong

Tipo	Descrição
Backport	HIVE-20868: O SMB Join falha de forma intermitente quando um filho está fazendo login TezDummyOperator getFinalOp MapRecord Processor
Backport	HIVE-20853: Exponha ShuffleHandler .register Dag na API do daemon llap
Backport	HIVE-20850: enviar o caso condicional das projeções para as tabelas de dimensões, se possível
Backport	HIVE-20842: corrigir a lógica introduzida no HIVE-20660 para estimar estatísticas para agrupamento
Backport	HIVE-20839: erro “Cannot find field” durante a junção de hash particionada dinamicamente
Backport	HIVE-20835: a interação entre restrições e regravação de MV pode criar um loop no planejador do Calcite
Backport	HIVE-20834: QueryResultCache Entradas do Hive que mantêm referência a partir da consulta em cache SemanticAnalyzer
Backport	HIVE-20830: falha na declaração JdbcStora geHandler de consulta de intervalo em alguns casos
Backport	HIVE-20829: JdbcStorageHandler range split lança NPE
Backport	HIVE-20827: resultados inconsistentes para matrizes vazias

Tipo	Descrição
Backport	HIVE-20826: HiveSemiJoin Regra aprimorada para converter junção + grupo do lado esquerdo para semijunção esquerda
Backport	HIVE-20821: reescrever SUM0 na combinação SUM + COALESCE
Backport	HIVE-20815: JdbcRecordReader .next não deve comer exceção
Backport	HIVE-20813: udf to_epoch_milli também precisa dar suporte a carimbo de data/hora sem fuso horário.
Backport	HIVE-20804: melhorias adicionais no agrupamento por otimização com restrições
Backport	HIVE-20792: inserir carimbo e data/hora com zonas trunca os dados
Backport	HIVE-20788: a redução estendida de SJ pode retroceder colunas incorretamente ao serem criados filtros
Backport	HIVE-20778: a reordenação de junções pode não ser acionada se todas as junções no plano são criadas pela lógica de decorrelação
Backport	HIVE-20772: registrar contadores de CPU por tarefa no LLAP
Backport	HIVE-20768: adicionar UDF de janela em cascata
Backport	HIVE-20767: vários projetos entre operadores de junção podem afetar a reordenação da junção usando restrições

Tipo	Descrição
Backport	HIVE-20762: o intervalo de limpeza do NOTIFICATION_LOG está embutido em código como 60 s e é muito pequeno
Backport	HIVE-20761: selecionar para atualização na tabela notification_sequence tem intervalo de repetição e contagem de tentativas muito pequena
Backport	HIVE-20751: atualizar a versão do Arrow para 0.10.0
Backport	HIVE-20746: HiveProtoHookLogger não fecha o arquivo no final do dia.
Backport	HIVE-20744: usar restrições SQL para melhorar o algoritmo de reordenação de junções
Backport	HIVE-20740: Remova o bloqueio global no método.setConf. ObjectStore Este cherrypick envia para backport o HIVE-20740 destinado ao Hive 3.2 e 4.x a 3.1.x
Backport	HIVE-20734: Beeline: quando o beeline-s ite.xml está no beeline e a CLI do hive redirecio na para o beeline, ele deve usar o nome de usuário/senha fictícia do sistema em vez de solicitar um
Backport	HIVE-20731: o arquivo de armazenamento de chaves deve ser autorizado JdbcStora geHandler
Backport	HIVE-20720: adicionar opção de coluna de partição ao manipulador JDBC

Tipo	Descrição
Backport	HIVE-20719: a instrução SELECT apresenta falha depois de UPDATE com otimização e vetorização de hive.optimize.sort.dynamic. partition ativadas
Backport	HIVE-20718: adicionar o driver perf cli com restrições
Backport	HIVE-20716: definir o valor padrão de hive.cbo. stats.correlated.multi.key.joins como verdadeiro
Backport	HIVE-20712: HivePointLookupOptimizer deve extrair casos profundos
Backport	HIVE-20710: o dobramento constante pode não criar constantes nulas sem tipos
Backport	HIVE-20706: external_jdbc_table2.q falhando intermitentemente
Backport	HIVE-20704: Estenda HivePreFilteringRule para suportar outras funções
Backport	HIVE-20703: colocar a otimização dinâmica da partição de classificação em uma decisão baseada em custos
Backport	HIVE-20702: considerar a sobrecarga das estimativas conscientes da estrutura de dados durante a seleção da junção do mapa
Backport	HIVE-20692: habilitar o dobramento de expressões NOT x IS (NOT) [TRUE|FALSE]
Backport	HIVE-20691: Corrija org.apache.hadoop. hive.cli. TestMiniLlapCliDriver. testCliDriver[cttl]

Tipo	Descrição
Backport	HIVE-20682: a execução de consultas assíncronas pode falhar se a sessionHive compartilhada for fechada pelo thread principal
Backport	HIVE-20676: HiveServer 2: não Privilege Synchronizer está definido para o status daemon
Backport	HIVE-20660: a estimativa de agrupamento por estatísticas pode ser melhorada ao ser limitado o número total de linhas à tabela de origem
Backport	HIVE-20652: junção JdbcStorageHandler push de duas fontes de dados diferentes ao driver jdbc
Backport	HIVE-20651: JdbcStorageHandler a senha deve ser criptografada
Backport	HIVE-20649: gerenciador de memória com reconhecimento de LLAP para gravadores do Orc
Backport	HIVE-20648: LLAP: o grupo de vetores por operador deve usar memória por executor
Backport	HIVE-20646: a condição do filtro de partição não é enviada para a consulta do metastore se ela tem IS NOT NULL
Backport	HIVE-20644: evitar expor informações confidenciais por meio de uma exceção de Runtime do Hive
Backport	HIVE-20636: melhorar a estimativa do número de valores nulos após a junção externa

Tipo	Descrição
Backport	HIVE-20632: a consulta com UDF de get_splits apresenta falha se a visão materializada é criada na tabela consultada
Backport	HIVE-20627: Consultas assíncronas simultâneas falham intermitentemente e causam vazamento de memória LockException
Backport	HIVE-20623: trabalho compartilhado: estender o compartilhamento de entradas de cache mapa-junção do LLAP
Backport	HIVE-20619: Incluir MultiDelimitSerDe em 2 por padrão HiveServer
Backport	HIVE-20618: Durante a seleção de junções, BucketMapJoin pode ser escolhido para tabelas sem compartimentos
Backport	HIVE-20617: corrigir o tipo de constantes nas expressões IN para ter o tipo correto
Backport	HIVE-20612: criar um novo sinalizador de correlação de várias chaves de junção para CBO
Backport	HIVE-20603: erro “Wrong FS” na inserção na partição depois da alteração do sistema de arquivos de localização da tabela
Backport	HIVE-20601: EnvironmentContext nulo no evento ALTER_PARTITION em DbNotificationListener
Backport	HIVE-20583: Use o nome de host canônico somente para autenticação kerberos em HiveConnection

Tipo	Descrição
Backport	HIVE-20582: tornar o hflush no registro em log de protótipos do Hive configurável
Backport	HIVE-20563: vetorização: a expressão CASE WHEN falha quando o tipo THEN/ELSE e o tipo de resultado são diferentes
Backport	HIVE-20558: alterar o padrão de hive.hash table.key.count.adjustment para 0,99
Backport	HIVE-20552: Obtenha o esquema de forma mais rápida LogicalPlan
Backport	HIVE-20550: Alterne a Web HCat para usar o beeline para enviar consultas do Hive
Backport	HIVE-20537: estimativas de junções de várias colunas com colunas não correlacionadas diferentes no CBO e no Hive
Backport	HIVE-20524: a verificação da evolução do esquema foi danificada na passagem da versão 2 do Hive para a versão 3 para ALTER TABLE VARCHAR para DECIMAL
Backport	HIVE-20522: HiveFilterSetOpTransposeRul e pode gerar erro de declaração devido à nulidade dos campos
Backport	HIVE-20521: HS2 doas=True tem problema de permissão com hadoop.tmp.dir, com sistema de arquivos MR e S3A

Tipo	Descrição
Backport	HIVE-20515: esvaziar os resultados da consulta ao usar o cache de resultados e o diretório temporário da consulta, o diretório do cache de resultados em diferentes sistemas de arquivos
Backport	HIVE-20508: o Hive não é compatível com nomes de usuário do tipo “user@realm”
Backport	HIVE-20507: Beeline: adicione um comando utilitário para recuperar todos os uris do beeline-site.xml
Backport	HIVE-20505: atualizar org.openjdk.jmh:jmh-core para 1.21
Backport	HIVE-20503: usar estimativas conscientes da estrutura de dados durante a seleção da junção do mapa
Backport	HIVE-20498: suporte ao tipo de data para coleta automática de estatísticas de colunas
Backport	HIVE-20496: Vetorização: PTF vetorizado IllegalStateException
Backport	HIVE-20494: O genérico UDFRestrict Informati onSchema está quebrado após o HIVE-19440
Backport	HIVE-20477: não OptimizedSql é mostrado se a expressão contiver INs
Backport	HIVE-20467: Permitir se não EXISTS/IF EXISTS in Resource plan creation/drop
Backport	HIVE-20462: “CREATE VIEW IF NOT EXISTS” falha se a visão já existe

Tipo	Descrição
Backport	HIVE-20455: Registre a velocidade do arquivo security.authorization. PrivilegeSynchroniz er.correr
Backport	HIVE-20439: usar o limite de memória inflado durante a seleção da junção para llap
Backport	HIVE-20433: a conversão implícita de string em carimbo de data/hora é lenta
Backport	HIVE-20432: regravar BETWEEN em IN para tipos de inteiros para estimativa de estatísticas
Backport	HIVE-20423: definir NULLS LAST como a ordem nula padrão
Backport	HIVE-20418: a E/S do LLAP IO pode não manipular arquivos ORC que tenham índice de linha desabilitado corretamente para consultas sem colunas selecionadas
Backport	HIVE-20412: NPE em HiveMetaHook
Backport	HIVE-20406: aglutinação aninhada fornecendo resultados incorretos
Backport	HIVE-20399: CTAS com um local de tabela personalizado que não é totalmente qualificado falha para tabelas MM
Backport	HIVE-20393: Redução de semijunção: FordPP se comporta de forma inconsistente markSemiJ oin
Backport	HIVE-20391: HiveAggregateReduceFunction sRule pode inferir tipo de retorno errado ao decompor a função agregada

Tipo	Descrição
Backport	HIVE-20383: problemas de nome de fila inválido e de sincronização no hook de eventos protocolo do hive.
Backport	HIVE-20367: vetorização: suporte ao streaming para PTF AVG, MAX, MIN, SUM
Backport	HIVE-20366: as estimativas das estatísticas de query78 de TPC-DS estão desativadas para o filtro nulo
Backport	HIVE-20364: atualizar o padrão para hive.map.aggr.hash.min.reduction
Backport	HIVE-20352: vetorização: suporte à função de agrupamento
Backport	HIVE-20347: hive.optimize.sort.dynamic.partition deve funcionar com CTAS e MV particionados
Backport	HIVE-20345: o descarte de dados pode travar se as tabelas forem excluídas em uma chamada diferente
Backport	HIVE-20343: Hive 3: CTAS não respeita transactional_properties
Backport	HIVE-20340: O Druid precisa de explícito de timestamp para STRING quando CASTs a saída da função timestamp é usada como Strin
Backport	HIVE-20339: vetorização: suspender restriçõe s desnecessárias que fazem com que alguns PTFs com RANK não sejam vetorizados

Tipo	Descrição
Backport	HIVE-20337: CachedStore: getPartitionsBy Expr não está preenchendo a lista de partições corretamente
Backport	HIVE-20336: políticas de mascaramento e filtragem para visões materializadas
Backport	HIVE-20326: criar restrições com RELY como padrão em vez de NO RELY
Backport	HIVE-20321: Vetorização: reduza o tamanho da memória de 1 coluna para <1 VectorHas hKeyWrapper CacheLine
Backport	HIVE-20320: ativar o sinalizador hive.opti mize.remove.sq_count_check
Backport	HIVE-20315: vetorização: corrigir mais problemas de NULL/Resultados incorretos e evitar conversões desnecessárias
Backport	HIVE-20314: incluir a remoção de partições na regravação da visão materializada
Backport	HIVE-20312: Permita que os clientes do Arrow usem seus próprios com BufferAllocator LlapOutputFormatService
Backport	HIVE-20302: LLAP: a execução não vetorizad a em E/S ignora colunas virtuais, incluindo ROW__ID
Backport	HIVE-20300: VectorFileSinkArrowOperator
Backport	HIVE-20299: corrida potencial no teste unitário de assinante do LLAP

Tipo	Descrição
Backport	HIVE-20296: Melhore HivePointLookupOptimizerRule para poder extrair de contextos mais sofisticados
Backport	HIVE-20294: vetorização: corrigir problemas de NULL/Resultados incorretos em COALESCE/ELT
Backport	HIVE-20292: ordem de junção incorreta em query93 de tpcds com restrição primária definida
Backport	HIVE-20290: Inicialização lenta para que não ArrowColumnarBatchSerDe aloque buffers durante GetSplits
Backport	HIVE-20281: SharedWorkOptimizer falha com 'o conteúdo do cache do operador e a diferença real do plano'
Backport	HIVE-20277: vetorização: expressões de caso que retornam BOOLEAN não são compatíveis com FILTER
Backport	HIVE-20267: expandir WebUI para incluir um formulário para configurar dinamicamente os níveis de log
Backport	HIVE-20263: Erro de digitação na variável HiveReduceExpressionsWithStatsRule
Backport	HIVE-20260: o NDV de uma coluna não deve ser escalado quando a contagem de linhas é alterada pelo filtro em outra coluna

Tipo	Descrição
Backport	HIVE-20252: redução de semijunção: os ciclos devido à ramificação da semijunção podem permanecer sem ser detectados se um lado pequeno da tabela tem uma junção de mapa upstream.
Backport	HIVE-20245: vetorização: corrigir problemas de NULL/Resultados incorretos em BETWEEN/IN
Backport	HIVE-20241: suporte a especificação do particionamento em instruções CTAS
Backport	HIVE-20240: redução de semijunção: usar a variável local para verificar a condição da tabela externa
Backport	HIVE-20226: O HMS getNextNotification lançará uma exceção quando a solicitação MaxEvents exceder as max_rows da tabela
Backport	HIVE-20225: SerDe para suportar o formato binário Teradata
Backport	HIVE-20213: atualizar o Calcite para 1.17.0
Backport	HIVE-20212: Hiveserver2 no modo http emitindo a métrica default.General.op en_connections incorretamente
Backport	HIVE-20210: O otimizador Simple Fetch deve fazer com que o filtro na coluna sem partição e MapReduce a conversão sejam mínimas
Backport	HIVE-20209: a conexão do Metastore falha na primeira tentativa no despejo da replicação

Tipo	Descrição
Backport	HIVE-20207: vetorização: corrigir problemas de NULL/Resultados incorretos em Filtrar/Comparar
Backport	HIVE-20204: conversão de tipo durante IN
Backport	HIVE-20203: A seta vaza um SerDe DirectByteBuffer
Backport	HIVE-20197: Vetorização: adicione o teste DECIMAL_64, adicione aritmética e adicione mais Date/Interval/Timestamp agregação GROUP BY
Backport	HIVE-20193: cboInfo não está presente no json do plano de explicação
Backport	HIVE-20192: HS2 com o metastore incorporado está vazando objetos do Manager JDOPersistence
Backport	HIVE-20183: a inserção de uma tabela com bucket pode causar perda de dados, se a tabela de origem contém um bucket vazio
Backport	HIVE-20177: Vetorização: reduza a alocação no modo Streaming KeyWrapper GroupBy
Backport	HIVE-20174: vetorização: corrigir problemas de NULL/Resultados incorretos nas funções de agregação GROUP BY
Backport	HIVE-20172: StatsUpdater falha com a exceção GSS ao tentar se conectar ao metastore remoto

Tipo	Descrição
Backport	HIVE-20153: UDF Count and Sum consome mais memória no Hive 2+
Backport	HIVE-20152: redefinir o estado do banco de dados, quando o repl dump falha, para que a renomeação da tabela possa ser feita
Backport	HIVE-20149: falha/tempo limite TestHiveCli
Backport	HIVE-20130: melhor registro em log para sincronizador de esquemas de informações
Backport	HIVE-20129: reverter para a evolução do esquema baseada em posição para tabelas do orc
Backport	HIVE-2018:. SessionStateUserAuthenticator getGroupNames
Backport	HIVE-20116: TezTask está usando o registrador principal
Backport	HIVE-20115: as tabelas Acid não devem usar a verificação de rodapé para analisar
Backport	HIVE-20103: WM: somente contadores de DAGs agregados se pelo menos um é usado
Backport	HIVE-20101: BloomKFilter: Evite usar inteiramente as matrizes locais de bytes []
Backport	HIVE-20100: OpTraits : Select Optraits deve parar quando uma incompatibilidade for detectada
Backport	HIVE-20098: estatísticas: NPE na obtenção de estatísticas de partição da coluna Data

Tipo	Descrição
Backport	HIVE-20095: corrigir o atributo para enviar computação para tabelas externas do jdbc
Backport	HIVE-20093: LlapOutputFomatService: Use ArrowBuf com Netty para contabilidade
Backport	HIVE-20090: estender a criação de filtros de redução de semijunção para possibilitar a descoberta de novas oportunidades
Backport	HIVE-20088: o caminho do local da configura ção do Beeline está montado incorretamente
Backport	HIVE-20082: HiveDecimal a conversão para string não formata o decimal corretamente
Backport	HIVE-20069: corrigir a reotimização em caso de otimização de DPP e semijunção
Backport	HIVE-20051: ignorar autorização para tabelas temporárias
Backport	HIVE-20044: Serde do Arrow deve preencher valores de caracteres e manipular cadeias vazias corretamente
Backport	HIVE-20028: a configuração de cache do cliente do Metastore é usada incorretamente
Backport	HIVE-20025: Limpeza de arquivos de eventos criados por HiveProtoLoggingHook
Backport	HIVE-20020: o arquivo jar de contribuição do Hive não deve estar na biblioteca
Backport	HIVE-20013: adicionar uma conversão implícita ao tipo de data para a função to_date

Tipo	Descrição
Backport	HIVE-20011: afastar-se do modo de anexação no hook de registro em log de protocolo
Backport	HIVE-20005: acid_table_stats, acid_no_buckets etc -alteração do resultado da consulta na ramificação
Backport	HIVE-20004: Escala errada usada por ConvertDecimal 64 ToDecimal resulta em resultados incorretos
Backport	HIVE-19995: agregar tráfego de linhas para tabelas acid
Backport	HIVE-19993: não é possível usar um alias de tabela que também aparece como nome de coluna
Backport	HIVE-19992: Vetorização: continuação do HIVE-19951 --> adicionar chamada a. SchemaEvolution isOnlyImplicitConversão para desativar a E/S LLAP codificada para ORC somente quando a conversão do tipo de dados não estiver implícita
Backport	HIVE-19989: O Metastore usa um nome de aplicativo errado para métricas HADOOP2
Backport	HIVE-19981: As tabelas gerenciadas convertid as em tabelas externas pelo HiveStric tManagedMigration utilitário devem ser configuradas para excluir dados quando a tabela for descartada
Backport	HIVE-19967: SMB Join: Need Optraits for ala GBY Op PTFOperator

Tipo	Descrição
Backport	HIVE-19935: sessão do Hive WM encerrada : falha ao atualizar a contagem de tarefas do LLAP
Backport	HIVE-19924: tarefas de Tag distcp executadas por Repl Load
Backport	HIVE-19891: inserir em tabelas externas com diretórios de partições personalizados pode causar perda de dados
Backport	HIVE-19850: a remoção de partições dinâmicas no Tez está levando ao erro “No work found for tablescan”
Backport	HIVE-19806: classificar a saída de qtests para evitar estranheza nos resultados de teste
Backport	HIVE-19770: suporte para CBO para consultas com várias colunas iguais na seleção
Backport	HIVE-19769: criar objetos dedicados para nomes de banco de dados e tabelas
Backport	HIVE-19765: Adicione testes específicos do Parquet ao BlobstoreCliDriver
Backport	HIVE-19759: Teste escamoso: # TestRpc testServerPort
Backport	HIVE-19711: ferramenta de refatorar no esquema do Hive
Backport	HIVE-19701: getDelegationToken FromMetaS tore não precisa ser sincronizado

Tipo	Descrição
Backport	HIVE-19694: a instrução Criar visão materializada (MV) deve verificar se há conflitos de nome da MV antes de executar a instrução SQL da MV.
Backport	HIVE-19674: grupo por constantes decimais enviado para uma tabela do Druid
Backport	HIVE-19668: Mais de 30% da pilha desperdiçada pela duplicação do org.antlr.runtime.CommonToken's e sequências duplicadas
Backport	HIVE-19663: refatorar a geração de relatórios de E/S do LLAP
Backport	HIVE-19661: alterne o Hive UDFs para usar o mecanismo regex Re2J
Backport	HIVE-19628: possível NPE no TestSigning do LLAP
Backport	HIVE-19568: HS2 HA ativo/passivo: não permitir conexão direta com a instância passiva HS2
Backport	HIVE-19564: vetorização: corrigir problemas de NULL/Resultados incorretos em Arithmetic
Backport	HIVE-19552: Habilite #druidkafkamini_basic .q TestMiniDruidKafkaCliDriver
Backport	HIVE-19432: GetTablesOperation é muito lento se o hive tiver muitos bancos de dados e tabelas
Backport	HIVE-19360: CBO: Adicione um “SQL otimizado” ao objeto QueryPlan

Tipo	Descrição
Backport	HIVE-19326: coleta automática de estatísticas: agregação incorreta durante consultas UNION
Backport	HIVE-19313: os testes estão falhando TestJdbcWith DBToken StoreNoDoAs
Backport	HIVE-19285: Adicione registros às subclasses de MetaDataOperation
Backport	HIVE-19235: atualizar arquivos dourados para testes Minimr
Backport	HIVE-19104: Quando o teste MetaStore é iniciado com uma nova tentativa, as instâncias devem ser independentes
Backport	HIVE-18986: A renomeação da tabela executará java.lang. StackOverflowError em DataNucleus se a tabela contiver um grande número de colunas
Backport	HIVE-18920: CBO: inicializar os provedores Janino antes da primeira consulta
Backport	HIVE-18873: Ignorar a pressão silenciosa do predicado para MR pode fazer com que os manipuladores de armazenamento produzam resultados HiveInputFormat errôneos
Backport	HIVE-18871: erro de execução do hive no tez devido à definição de hive.aux.jars.path como hdfs://
Backport	HIVE-18725: melhorar o tratamento de erros para subconsultas se houver uma referência de coluna errada

Tipo	Descrição
Backport	HIVE-18696: As pastas de partições podem não ser limpas adequadamente no método.add_partitions_core se um HiveMetaStore
Backport	HIVE-18453: ACID: adicionar a sintaxe “CREATE TRANSACTIONAL TABLE” para unificar o suporte a ACID ORC e Parquet
Backport	HIVE-18201: desabilitar XPROD_EDGE para sq_count_chec
Backport	HIVE-18140: as estatísticas de tabelas particionadas podem dar errado em estatísticas básicas em maiúsculas e minúsculas
Backport	HIVE-17921: a agregação com estrutura no LLAP produz um resultado incorreto
Backport	HIVE-17896: TopNKey: Crie um operador superior vetorizável autônomo NKey
Backport	HIVE-17840: HiveMetaStore come exceção se TransactionalListeners.notifyEvent falhar
Backport	HIVE-17043: remover colunas não exclusivas do grupo por chaves se não são referenciadas posteriormente
Backport	HIVE-17040: eliminação de junções na presença de relação FK
Backport	HIVE-16839: chamadas desequilibradas para openTransaction/commitTransaction ao alterar a mesma partição simultaneamente

Tipo	Descrição
Backport	HIVE-16100: o otimizador de partições dinâmicas classificadas perde operadores irmãos
Backport	HIVE-15956: StackOverflowError quando descartar muitas partições
Backport	HIVE-15177: a autenticação com o hive falha quando o tipo de autenticação do kerberos é definido como fromSubject e a entidade principal contém _HOST
Backport	HIVE-14898: não HS2 deve registrar a pilha de chamadas devido a um erro de cabeçalho de autenticação vazio
Backport	HIVE-14493: suporte de particionamento para visualizações materializadas
Backport	HIVE-14431: reconhecer COALESCE como CASE
Backport	HIVE-13457: Crie pontos de extremidade da API HS2 REST para monitorar informações
Backport	HIVE-12342: definir o valor padrão de hive.optimize.index.filter como verdadeiro
Backport	HIVE-10296: exceção de conversão observada quando o hive executa uma consulta de várias junções no metastore
Backport	HIVE-6980: descartar a tabela usando sql direto

Amazon EMR 6.6.0 - Alterações na configuração do Hive

- Como parte da alteração do OSS, o HIVE-20703, a propriedade para classificar partições dinâmicas, `hive.optimize.sort.dynamic.partition`, foi substituída por `hive.optimize.sort.dynamic.partition.threshold`.

  A configuração `hive.optimize.sort.dynamic.partition.threshold` tem os seguintes valores potenciais:

Valor	Descrição
`0` (padrão)	Torna a otimização para classificar partições dinâmicas uma decisão baseada em custos quando arquivos ORC são usados. O número máximo de gravadores permitidos nas consultas INSERT é calculado com base em (memória do executor/contêiner) * (porcentagem da memória ocupada pelo orc) dividido pela memória máxima (tamanho da faixa) ocupada por um único gravador.
`-1`	Desabilita a otimização para classificar partições dinâmicas por completo.
`1`	Permite a classificação global de partições dinâmicas. Isso mantém apenas um gravador de registros aberto para cada valor de partição no redutor, reduzindo, portanto, a pressão de memória nos redutores.
`2` (ou número inteiro maior)	Diz ao Hive para usar o número inteiro especificado como limite para o número máximo de gravadores.

Amazon EMR 6.6.0 - Problemas conhecidos do Hive

- Consultas com funções de janela na mesma coluna da junção podem levar a transformações inválidas, conforme informado no HIVE-25278, e causar resultados incorretos ou falhas nas consultas. Como solução alternativa, desabilite o CBO em nível de consulta para essas consultas. Entre em contato com o AWS suporte para obter mais informações.
- O Amazon EMR 6.6.0 inclui o software Hive versão 3.1.2. O Hive 3.1.2 introduz um atributo que divide arquivos de texto se eles contêm um cabeçalho e um rodapé (HIVE-21924). O Apache Tez App Master lê cada um dos seus arquivos para determinar pontos de deslocamento no intervalo de

dados. Esses comportamentos combinados podem afetar negativamente a performance se suas consultas leem um grande número de pequenos arquivos de texto. Como solução alternativa, use `CombineHiveInputFormat` e ajuste o tamanho máximo da divisão configurando as seguintes propriedades:

```
SET hive.tez.input.format=org.apache.hadoop.hive.ql.io.CombineHiveInputFormat;
SET mapreduce.input.fileinputformat.split.maxsize=16777216;
```

- Com o Amazon EMR 6.6.0 a 6.9.x, as consultas INSERT com partição dinâmica e uma cláusula ORDER BY ou SORT BY sempre terá dois redutores. Esse problema é causado pela alteração do OSS HIVE-20703, que coloca a otimização da partição dinâmica de classificação sob uma decisão baseada em custos. Se sua workload não exigir a classificação de partições dinâmicas, recomendamos que você defina a propriedade `hive.optimize.sort.dynamic.partition.threshold` como `-1` para desabilitar o novo atributo e obter o número de redutores calculado corretamente. Esse problema foi corrigido no OSS Hive como parte do HIVE-22269 e foi corrigido no Amazon EMR 6.10.0.

Hudi

O Apache Hudi é uma estrutura de gerenciamento de dados de código aberto usada para simplificar o processamento incremental de dados e o desenvolvimento de pipelines de dados ao fornecer recursos de inserção, atualização, upsert e exclusão em nível de registro. Upsert refere-se à capacidade de inserir registros em um conjunto de dados existente se eles ainda não existirem ou atualizá-los se existirem. Ao gerenciar eficientemente a forma como os dados são apresentados no Amazon S3, o Hudi permite que os dados sejam ingeridos e atualizados quase em tempo real. O Hudi mantém cuidadosamente metadados das ações realizadas no conjunto de dados para ajudar a garantir que as ações sejam atômicas e consistentes.

O Hudi é integrado ao Apache Spark, ao Apache Hive e ao Presto. Nas versões 6.1.0 e posteriores do Amazon EMR, o Hudi também está integrado ao Trino (PrestoSQL).

Nas versões 5.28.0 e posteriores, o Amazon EMR instala os componentes do Hudi por padrão quando o Spark, o Hive, o Presto ou o Flink está instalado. Você pode usar o Spark ou o DeltaStreamer utilitário Hudi para criar ou atualizar conjuntos de dados Hudi. É possível usar o Hive, o Spark, o Presto ou o Flink para consultar um conjunto de dados do Hudi interativamente ou criar pipelines de processamento de dados usando extração incremental. Extração incremental refere-se à capacidade de extrair apenas os dados que foram alterados entre duas ações.

Esses atributos tornam o Hudi adequado para os seguintes casos de uso:

- Trabalhar com dados de streaming de sensores e outros dispositivos da Internet das Coisas (IoT) que exigem eventos específicos de inserção e atualização de dados.
- Cumprir os regulamentos de privacidade de dados em aplicativos em que os usuários possam optar por serem esquecidos ou modificar seu consentimento para a forma como os dados podem ser utilizados.
- Implementar um sistema de captura de dados de alteração (CDC) que permita aplicar alterações a um conjunto de dados ao longo do tempo.

A tabela a seguir lista a versão do Hudi incluída na versão mais recente da série 7.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Hudi.

Para obter a versão dos componentes instalados com o Hudi nesse lançamento, consulte Release 7.9.0 Component Versions.

Informações de versão do Hudi para o emr-7.9.0

Rótulo de versão do Amazon EMR	Versão do Hudi	Componentes instalados com o Hudi
emr-7.9.0	Hudi 0.15.0-amzn-6	Not available.

A tabela a seguir lista a versão do Hudi incluída na versão mais recente da série 6.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Hudi.

Para obter a versão dos componentes instalados com o Hudi nessa versão, consulte Release 6.15.0 Component Versions.

Informações de versão do Hudi para o emr-6.15.0

Rótulo de versão do Amazon EMR	Versão do Hudi	Componentes instalados com o Hudi
emr-6.15.0	Hudi 0.14.0-amzn-0	Not available.

> **Note**
>
> A versão 6.8.0 do Amazon EMR vem com o Apache Hudi 0.11.1; no entanto, os clusters do Amazon EMR 6.8.0 também são compatíveis com o código aberto `hudi-spark3.3-bundle_2.12` do Hudi 0.12.0.

A tabela a seguir lista a versão do Hudi incluída na versão mais recente da série 5.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Hudi.

Para obter a versão dos componentes instalados com o Hudi nesse lançamento, consulte Release 5.36.2 Component Versions.

Informações de versão do Hudi para o emr-5.36.2

Rótulo de versão do Amazon EMR	Versão do Hudi	Componentes instalados com o Hudi
emr-5.36.2	Hudi 0.10.1-amzn-1	Not available.

Tópicos

- Como o Hudi funciona
- Considerações e limitações para usar o Hudi no Amazon EMR
- Criar um cluster com o Hudi instalado
- Trabalhe com um conjunto de dados do Hudi
- Usar a CLI do Hudi
- Histórico de versões do Hudi

Como o Hudi funciona

Ao usar o Hudi com o Amazon EMR, você pode gravar dados no conjunto de dados usando a API da fonte de dados do Spark ou o utilitário Hudi. DeltaStreamer O Hudi organiza um conjunto de dados em uma estrutura de diretórios particionada em um *`basepath`* que é semelhante a uma tabela tradicional do Hive. As especificidades de como os dados são definidos como arquivos nesses diretórios dependem do tipo de conjunto de dados escolhido. É possível escolher Copiar na gravação (CoW – Copy on Write) ou Mesclar na leitura (MoR – Merge on Read).

Independentemente do tipo de conjunto de dados, cada partição em um conjunto de dados é identificada exclusivamente por seu `partitionpath` relativo ao `basepath`. Dentro de cada partição, os registros são distribuídos em vários arquivos de dados. Para obter mais informações, consulte File management na documentação do Apache Hudi.

Cada ação no Hudi tem uma confirmação correspondente, identificada por um carimbo de data/hora monotonicamente crescente conhecido como Instant. O Hudi mantém uma série de todas as ações executadas no conjunto de dados como uma linha do tempo. O Hudi conta com a linha do tempo para fornecer isolamento de snapshot entre leitores e gravadores e para permitir a reversão para um momento anterior específico. Para obter mais informações sobre as ações registradas pelo Hudi e o estado das ações, consulte Linha do tempo na documentação do Apache Hudi.

Noções básicas sobre os tipos de armazenamento de conjuntos de dados: copiar na gravação ou mesclar na leitura

Ao criar um conjunto de dados do Hudi, você especifica se o conjunto de dados é do tipo copiar na gravação ou mesclar na leitura.

- Copy on Write (CoW – Copiar na gravação): os dados são armazenados em um formato colunar (Parquet), e cada atualização cria uma nova versão dos arquivos durante uma gravação. CoW é o tipo de armazenamento padrão.
- Merge on Read (MoR – Mesclar na leitura): os dados são armazenados usando uma combinação de formatos colunares (Parquet) e baseados em linha (Avro). As atualizações são registradas em arquivos delta baseados em linha e são compactadas conforme necessário para criar novas versões dos arquivos colunares.

Com conjuntos de dados CoW, sempre que há uma atualização para um registro, o arquivo que contém o registro é regravado com os valores atualizados. Com um conjunto de dados MoR, sempre que há uma atualização, o Hudi grava apenas a linha do registro alterado. MoR é mais adequado para cargas de trabalho com maior volume de gravações ou alterações e menor volume de leituras. CoW é mais adequado para cargas de trabalho com maior volume de leituras em dados que mudam com menos frequência.

O Hudi fornece três visualizações lógicas para acessar os dados:

- Visão otimizada para leitura: fornece o conjunto de dados confirmado mais recente de tabelas CoW e o conjunto de dados compactados mais recente de tabelas MoR.
- Visão incremental: fornece um fluxo de alterações entre duas ações de um conjunto de dados CoW para alimentar trabalhos downstream e extrair, transformar e carregar (ETL) fluxos de trabalho.
- Visão em tempo real: fornece os dados confirmados mais recentes de uma tabela MoR ao mesclar os arquivos baseados em colunas e linhas em linha.

Quando você consulta a visualização otimizada para leitura, a consulta retorna todos os dados compactados, mas não inclui as últimas confirmações delta. A consulta desses dados apresenta bom desempenho de leitura, mas omite os dados mais recentes. Quando você consulta a visualização em tempo real, o Hudi mescla os dados compactados com as confirmações delta na leitura. Os dados mais recentes estão disponíveis para consulta, mas a sobrecarga computacional da mesclagem prejudica o desempenho da consulta. A capacidade de consultar dados compactados ou dados em tempo real permite a você escolher entre o desempenho e a flexibilidade nas consultas.

Para obter mais informações sobre as vantagens e desvantagens dos tipos de armazenamento, consulte Tipos e visões de armazenamento na documentação do Apache Hudi.

O Hudi cria duas tabelas no metastore do Hive para MoR: uma tabela com o nome especificado por você, que é uma visão otimizada para leitura, e uma tabela com o mesmo nome anexado com `_rt`, que é uma visão em tempo real. É possível consultar as duas tabelas.

Registrar um conjunto de dados do Hudi com seu metastore

Ao registrar uma tabela do Hudi com o metastore do Hive, você pode consultar tabelas do Hudi usando o Hive, o Spark SQL ou o Presto como faria com qualquer outra tabela. Além disso, é possível integrar o Hudi com o AWS Glue configurando o Hive e o Spark para usar o Catálogo de Dados do AWS Glue como metastore. Para tabelas MoR, o Hudi registra o conjunto de dados como duas tabelas no Metastore: uma tabela com o nome especificado por você, que é uma visão otimizada para leitura, e uma tabela com o mesmo nome anexado com `_rt`, que é uma visão em tempo real.

Você registra uma tabela do Hudi com o metastore do Hive ao usar o Spark para criar um conjunto de dados do Hudi ao definir a opção `HIVE_SYNC_ENABLED_OPT_KEY` como `"true"` e fornecendo outras propriedades necessárias. Para obter mais informações, consulte Trabalhe com um conjunto de dados do Hudi. Além disso, é possível usar o utilitário de linha de comando hive_sync_tool para registrar um conjunto de dados do Hudi como uma tabela no seu metastore, separadamente.

Considerações e limitações para usar o Hudi no Amazon EMR

- O campo de chave de registro não pode ser nulo nem estar em branco: o campo especificado como o campo de chave de registro não pode ter valores `null` nem em branco.
- Esquema atualizado por padrão na atualização e inserção — o Hudi fornece uma interface `HoodieRecordPayload` que determina como a entrada DataFrame e o conjunto de dados Hudi existente são mesclados para produzir um conjunto de dados novo e atualizado. O Hudi fornece uma implementação padrão dessa classe,`OverwriteWithLatestAvroPayload`, que substitui os registros existentes e atualiza o esquema conforme especificado na entrada. DataFrame Para personalizar essa lógica a fim de implementar atualizações parciais e de mesclagem, é possível fornecer uma implementação da interface `HoodieRecordPayload` usando o parâmetro `DataSourceWriteOptions.PAYLOAD_CLASS_OPT_KEY`.
- A exclusão requer um esquema: ao excluir, você precisa especificar a chave de registro, a chave de partição e os campos de chave de pré-combinação. Outras colunas podem ficar `null` ou em branco, mas o esquema completo é necessário.
- Limitações das tabelas MoR: as tabelas MoR não são compatíveis com pontos de salvamento. Você pode consultar tabelas MoR usando a visão otimizada para leitura ou a visão em tempo real

(*tableName*_rt) do Spark SQL, Presto ou Hive. O uso da visão otimizada para leitura expõe apenas os dados do arquivo básico e não expõe uma visão mesclada dos dados básicos e de log.

- Hive
  - Para registrar tabelas no metastore do Hive, o Hudi espera que o servidor Thrift do Hive esteja em execução na porta padrão `10000`. Se você substituir essa porta por uma porta personalizada, transmita a opção `HIVE_URL_OPT_KEY` conforme mostrado no exemplo a seguir.

    ```
    .option(DataSourceWriteOptions.HIVE_URL_OPT_KEY, "jdbc:hive2://localhost:override-port-number
    ```

  - O tipo de dados `timestamp` no Spark é registrado como tipo de dados `long` no Hive, e não como o tipo `timestamp` do Hive.
- Presto
  - O Presto não é compatível com a leitura de tabelas MoR em tempo real em versões do Hudi anteriores à 0.6.0.
  - O Presto só oferece suporte a consultas de snapshots.
  - Para o Presto interpretar corretamente colunas de conjuntos de dados do Hudi, defina o valor `hive.parquet_use_column_names` como `true`.
    - Para definir o valor de uma sessão, execute o seguinte comando no shell Presto:

      ```
      set session hive.parquet_use_column_names=true
      ```

    - Para definir o valor no nível do cluster, use a classificação de configuração `presto-connector-hive` para definir `hive.parquet.use_column_names` como `true`, conforme mostrado no exemplo a seguir. Para obter mais informações, consulte Configurar aplicações.

      ```
      [
        {
          "Classification": "presto-connector-hive",
          "Properties": {
            "hive.parquet.use-column-names": "true"
          }
        }
      ]
      ```

- HBase Índice

- A HBase versão usada para criar o Hudi pode ser diferente da listada no Guia de versões do EMR. Para extrair as dependências corretas da sua sessão do Spark, acesse o comando a seguir.

```
spark-shell \
--jars /usr/lib/spark/external/lib/spark-avro.jar,/usr/lib/hudi/cli/lib/*.jar \
--conf "spark.serializer=org.apache.spark.serializer.KryoSerializer" \
--conf "spark.sql.hive.convertMetastoreParquet=false"
```

- Configurações para melhor desempenho — Para o EMR 7.3+/Hudi 0.15+, recomenda-se que os clientes definam essa configuração para reduzir a sobrecarga de serialização do Kryo:

```
--conf 'spark.kryo.registrator=org.apache.spark.HoodieKryoRegistrar'
```

> **Note**
>
> Se você estiver usando controle de acesso refinado (FGAC) no EMR Serverless, essa configuração não é necessária, pois os usuários devem usar em vez de. JavaSerializer KryoSerializer

Criar um cluster com o Hudi instalado

Nas versões 5.28.0 e posteriores, o Amazon EMR instala os componentes do Hudi por padrão quando o Spark, o Hive, ou o Presto está instalado. Para usar o Hudi no Amazon EMR, crie um cluster com uma ou mais das seguintes aplicações instaladas:

- Hadoop
- Hive
- Spark
- Presto
- Flink

É possível criar um cluster usando o AWS Management Console AWS CLI, a ou a API do Amazon EMR.

Criar um cluster com o Hudi usando o AWS Management Console

1. Navegue até o novo console do Amazon EMR e selecione Alternar para o console antigo na navegação lateral. Para obter mais informações sobre o que esperar ao alternar para o console antigo, consulte Usar o console antigo.
2. Escolha Create cluster (Criar cluster), Go to advanced options (Ir para opções avançadas).
3. Em Software Configuration (Configuração de software), escolha emr-5.28.0 ou posterior para Release (Versão) e selecione Hadoop, Hive, Spark, Presto e Tez junto com outros aplicativos exigidos pelo seu cluster.
4. Configure outras opções conforme necessário para seu aplicativo e escolha Next (Avançar).
5. Configure as opções de Hardware e as General cluster settings (Configurações gerais do cluster) conforme desejado.
6. Para Opções de segurança, recomendamos que você selecione um par de EC2 chaves que possa ser usado para se conectar à linha de comando do nó principal usando SSH. Isso permite que você execute os comandos do shell do Spark, os comandos da CLI do Hive e os comandos da CLI do Hudi descritos neste guia.
7. Configure outras opções de segurança conforme desejado e escolha Create cluster (Criar cluster).

Trabalhe com um conjunto de dados do Hudi

O Hudi oferece suporte à inserção, atualização e exclusão de dados em conjuntos de dados do Hudi por meio do Spark. Para obter mais informações, consulte Gravar tabelas do Hudi na documentação do Apache Hudi.

Os exemplos a seguir demonstram como iniciar o shell interativo do Spark, usar o envio do Spark ou usar os Cadernos do Amazon EMR para trabalhar com o Hudi no Amazon EMR. Também é possível usar o DeltaStreamer utilitário Hudi ou outras ferramentas para gravação em um conjunto de dados. Ao longo desta seção, os exemplos demonstram como trabalhar com conjuntos de dados usando o shell do Spark durante a conexão com o nó principal usando SSH como usuário padrão `hadoop`.

Iniciar o shell do Spark usando as versões 6.7 e posteriores do Amazon EMR

Ao executar `spark-shell`, `spark-submit` ou `spark-sql` ao usar as versões 6.7.0 ou posteriores do Amazon EMR , passe os comandos a seguir.

> **Note**
>
> O Amazon EMR 6.7.0 usa o Apache Hudi 0.11.0-amzn-0, que contém melhorias significativas em relação às versões anteriores do Hudi. Para obter mais informações, consulte o Guia de migração do Apache Hudi 0.11.0. Os exemplos nesta guia refletem essas alterações.

Abrir o shell do Spark no nó primário

1. Conectar-se ao nó primário usando SSH. Para obter mais informações, consulte Connect to the primary node using SSH no Guia de gerenciamento do Amazon EMR.
2. Digite o seguinte comando para iniciar o shell do Spark. Para usar a PySpark concha, *`spark-shell`* substitua por*`pyspark`*.

```
spark-shell --jars /usr/lib/hudi/hudi-spark-bundle.jar \
--conf "spark.serializer=org.apache.spark.serializer.KryoSerializer" \
--conf
 "spark.sql.catalog.spark_catalog=org.apache.spark.sql.hudi.catalog.HoodieCatalog"
 \
--conf "spark.sql.extensions=org.apache.spark.sql.hudi.HoodieSparkSessionExtension"
```

Iniciar o shell do Spark usando as versões 6.6 e anteriores do Amazon EMR

Ao executar `spark-shell`, `spark-submit` ou `spark-sql` ao usar as versões 6.6.x ou anteriores do Amazon EMR , passe os comandos a seguir.

> **Note**
>
> - O Amazon EMR 6.2 e as versões 5.31 e posteriores (versões 0.6.x e posteriores do Hudi) podem omitir o `spark-avro.jar` da configuração.
> - O Amazon EMR 6.5 e as versões 5.35 e posteriores (versões 0.9.x e posteriores do Hudi) podem omitir o `spark.sql.hive.convertMetastoreParquet=false` da configuração.
> - O Amazon EMR 6.6 e as versões 5.36 e posteriores (versões 0.10.x e posteriores do Hudi) devem incluir a configuração `HoodieSparkSessionExtension` conforme descrito na Versão: Guia do Spark 0.10.0:

```
--conf
 "spark.sql.extensions=org.apache.spark.sql.hudi.HoodieSparkSessionExtension"
 \
```

Abrir o shell do Spark no nó primário

1. Conectar-se ao nó primário usando SSH. Para obter mais informações, consulte Connect to the primary node using SSH no Guia de gerenciamento do Amazon EMR.
2. Digite o seguinte comando para iniciar o shell do Spark. Para usar a PySpark concha, *spark-shell* substitua por*pyspark*.

```
spark-shell \
--conf "spark.serializer=org.apache.spark.serializer.KryoSerializer" \
--conf "spark.sql.hive.convertMetastoreParquet=false" \
--jars /usr/lib/hudi/hudi-spark-bundle.jar,/usr/lib/spark/external/lib/spark-
avro.jar
```

Usar o Hudi com os Cadernos do Amazon EMR usando as versões 6.7 e posteriores do Amazon EMR

Para usar o Hudi com os Cadernos do Amazon EMR, você deve primeiro copiar os arquivos jar do Hudi do sistema de arquivos local para o HDFS no nó principal do cluster do caderno. Em seguida, use o editor de caderno para configurar o caderno do EMR para usar o Hudi.

Usar o Hudi com os Cadernos do Amazon EMR

1. Crie e inicie um cluster para Cadernos do Amazon EMR. Para obter mais informações, consulte Creating Amazon EMR clusters for notebooks no Guia de gerenciamento do Amazon EMR.
2. Conecte-se ao nó principal do cluster usando SSH e copie os arquivos jar do sistema de arquivos local para o HDFS, conforme mostrado nos exemplos a seguir. No exemplo, criamos um diretório no HDFS para fins de clareza do gerenciamento de arquivos. É possível escolher seu próprio destino no HDFS, se desejar.

```
hdfs dfs -mkdir -p /apps/hudi/lib
```

```
hdfs dfs -copyFromLocal /usr/lib/hudi/hudi-spark-bundle.jar /apps/hudi/lib/hudi-
spark-bundle.jar
```

3. Abra o editor do notebook, insira o código do exemplo a seguir e execute-o.

```
%%configure
{ "conf": {
            "spark.jars":"hdfs:///apps/hudi/lib/hudi-spark-bundle.jar",
            "spark.serializer":"org.apache.spark.serializer.KryoSerializer",
            "spark.sql.catalog.spark_catalog":
 "org.apache.spark.sql.hudi.catalog.HoodieCatalog",

 "spark.sql.extensions":"org.apache.spark.sql.hudi.HoodieSparkSessionExtension"
          }}
```

Usar o Hudi com Cadernos do Amazon EMR usando as versões 6.6 e anteriores do Amazon EMR

Para usar o Hudi com os Cadernos do Amazon EMR, você deve primeiro copiar os arquivos jar do Hudi do sistema de arquivos local para o HDFS no nó principal do cluster do caderno. Em seguida, use o editor de caderno para configurar o caderno do EMR para usar o Hudi.

Usar o Hudi com os Cadernos do Amazon EMR

1. Crie e inicie um cluster para Cadernos do Amazon EMR. Para obter mais informações, consulte Creating Amazon EMR clusters for notebooks no Guia de gerenciamento do Amazon EMR.
2. Conecte-se ao nó principal do cluster usando SSH e copie os arquivos jar do sistema de arquivos local para o HDFS, conforme mostrado nos exemplos a seguir. No exemplo, criamos um diretório no HDFS para fins de clareza do gerenciamento de arquivos. É possível escolher seu próprio destino no HDFS, se desejar.

```
hdfs dfs -mkdir -p /apps/hudi/lib
```

```
hdfs dfs -copyFromLocal /usr/lib/hudi/hudi-spark-bundle.jar /apps/hudi/lib/hudi-
spark-bundle.jar
```

```
hdfs dfs -copyFromLocal /usr/lib/spark/external/lib/spark-avro.jar /apps/hudi/lib/
spark-avro.jar
```

3. Abra o editor do notebook, insira o código do exemplo a seguir e execute-o.

```
{ "conf": {
            "spark.jars":"hdfs:///apps/hudi/lib/hudi-spark-bundle.jar,hdfs:///apps/
hudi/lib/spark-avro.jar",
            "spark.serializer":"org.apache.spark.serializer.KryoSerializer",
            "spark.sql.hive.convertMetastoreParquet":"false"
          }}
```

Inicializar uma sessão do Spark para Hudi

Ao usar o Scala, você deve importar as seguintes classes na sessão do Spark. Isso precisa ser feito uma vez por sessão do Spark.

```
import org.apache.spark.sql.SaveMode
import org.apache.spark.sql.functions._
import org.apache.hudi.DataSourceWriteOptions
import org.apache.hudi.DataSourceReadOptions
import org.apache.hudi.config.HoodieWriteConfig
import org.apache.hudi.hive.MultiPartKeysValueExtractor
import org.apache.hudi.hive.HiveSyncConfig
import org.apache.hudi.sync.common.HoodieSyncConfig
```

Gravar em um conjunto de dados do Hudi

Os exemplos a seguir mostram como criar um DataFrame e gravá-lo como um conjunto de dados do Hudi.

> **Note**
>
> Para colar exemplos de código no shell do Spark, digite **`:paste`** no prompt, cole o exemplo e pressione **`CTRL + D`**.

Cada vez que você grava um DataFrame em um conjunto de dados Hudi, você deve especificar. `DataSourceWriteOptions` Muitas dessas opções provavelmente serão idênticas entre as

operações de gravação. O exemplo a seguir especifica opções comuns usando a variável *hudiOptions*, usada pelos exemplos subsequentes.

Gravar usando Scala com as versões 6.7 e posteriores do Amazon EMR

> **Note**
>
> O Amazon EMR 6.7.0 usa o Apache Hudi 0.11.0-amzn-0, que contém melhorias significativas em relação às versões anteriores do Hudi. Para obter mais informações, consulte o Guia de migração do Apache Hudi 0.11.0. Os exemplos nesta guia refletem essas alterações.

```
// Create a DataFrame
val inputDF = Seq(
 ("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
 ("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
 ("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
 ("103", "2015-01-01", "2015-01-01T13:51:40.519832Z"),
 ("104", "2015-01-02", "2015-01-01T12:15:00.512679Z"),
 ("105", "2015-01-02", "2015-01-01T13:51:42.248818Z")
 ).toDF("id", "creation_date", "last_update_time")

//Specify common DataSourceWriteOptions in the single hudiOptions variable
val hudiOptions = Map[String,String](
  HoodieWriteConfig.TBL_NAME.key -> "tableName",
  DataSourceWriteOptions.TABLE_TYPE.key -> "COPY_ON_WRITE",
  DataSourceWriteOptions.RECORDKEY_FIELD_OPT_KEY -> "id",
  DataSourceWriteOptions.PARTITIONPATH_FIELD_OPT_KEY -> "creation_date",
  DataSourceWriteOptions.PRECOMBINE_FIELD_OPT_KEY -> "last_update_time",
  DataSourceWriteOptions.HIVE_SYNC_ENABLED_OPT_KEY -> "true",
  DataSourceWriteOptions.HIVE_TABLE_OPT_KEY -> "tableName",
  DataSourceWriteOptions.HIVE_PARTITION_FIELDS_OPT_KEY -> "creation_date",
  HoodieSyncConfig.META_SYNC_PARTITION_EXTRACTOR_CLASS.key ->
 "org.apache.hudi.hive.MultiPartKeysValueExtractor",
  HoodieSyncConfig.META_SYNC_ENABLED.key -> "true",
  HiveSyncConfig.HIVE_SYNC_MODE.key -> "hms",
  HoodieSyncConfig.META_SYNC_TABLE_NAME.key -> "tableName",
  HoodieSyncConfig.META_SYNC_PARTITION_FIELDS.key -> "creation_date"
)

// Write the DataFrame as a Hudi dataset
(inputDF.write
```

```
    .format("hudi")
    .options(hudiOptions)
    .option(DataSourceWriteOptions.OPERATION_OPT_KEY,"insert")
    .mode(SaveMode.Overwrite)
    .save("s3://amzn-s3-demo-bucket/myhudidataset/"))
```

Gravar usando Scala com as versões 6.6 e anteriores do Amazon EMR

```
// Create a DataFrame
val inputDF = Seq(
 ("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
 ("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
 ("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
 ("103", "2015-01-01", "2015-01-01T13:51:40.519832Z"),
 ("104", "2015-01-02", "2015-01-01T12:15:00.512679Z"),
 ("105", "2015-01-02", "2015-01-01T13:51:42.248818Z")
 ).toDF("id", "creation_date", "last_update_time")

//Specify common DataSourceWriteOptions in the single hudiOptions variable
val hudiOptions = Map[String,String](
  HoodieWriteConfig.TABLE_NAME -> "tableName",
  DataSourceWriteOptions.TABLE_TYPE_OPT_KEY -> "COPY_ON_WRITE",
  DataSourceWriteOptions.RECORDKEY_FIELD_OPT_KEY -> "id",
  DataSourceWriteOptions.PARTITIONPATH_FIELD_OPT_KEY -> "creation_date",
  DataSourceWriteOptions.PRECOMBINE_FIELD_OPT_KEY -> "last_update_time",
  DataSourceWriteOptions.HIVE_SYNC_ENABLED_OPT_KEY -> "true",
  DataSourceWriteOptions.HIVE_TABLE_OPT_KEY -> "tableName",
  DataSourceWriteOptions.HIVE_PARTITION_FIELDS_OPT_KEY -> "creation_date",
  DataSourceWriteOptions.HIVE_PARTITION_EXTRACTOR_CLASS_OPT_KEY ->
 classOf[MultiPartKeysValueExtractor].getName
)

// Write the DataFrame as a Hudi dataset
(inputDF.write
    .format("org.apache.hudi")
    .option(DataSourceWriteOptions.OPERATION_OPT_KEY,
 DataSourceWriteOptions.INSERT_OPERATION_OPT_VAL)
    .options(hudiOptions)
    .mode(SaveMode.Overwrite)
    .save("s3://amzn-s3-demo-bucket/myhudidataset/"))
```

Escreva usando PySpark

```
# Create a DataFrame
inputDF = spark.createDataFrame(
    [
        ("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
        ("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
        ("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
        ("103", "2015-01-01", "2015-01-01T13:51:40.519832Z"),
        ("104", "2015-01-02", "2015-01-01T12:15:00.512679Z"),
        ("105", "2015-01-02", "2015-01-01T13:51:42.248818Z"),
    ],
    ["id", "creation_date", "last_update_time"]
)

# Specify common DataSourceWriteOptions in the single hudiOptions variable
hudiOptions = {
'hoodie.table.name': 'tableName',
'hoodie.datasource.write.recordkey.field': 'id',
'hoodie.datasource.write.partitionpath.field': 'creation_date',
'hoodie.datasource.write.precombine.field': 'last_update_time',
'hoodie.datasource.hive_sync.enable': 'true',
'hoodie.datasource.hive_sync.table': 'tableName',
'hoodie.datasource.hive_sync.partition_fields': 'creation_date',
'hoodie.datasource.hive_sync.partition_extractor_class':
 'org.apache.hudi.hive.MultiPartKeysValueExtractor'
}

# Write a DataFrame as a Hudi dataset
inputDF.write \
.format('org.apache.hudi') \
.option('hoodie.datasource.write.operation', 'insert') \
.options(**hudiOptions) \
.mode('overwrite') \
.save('s3://amzn-s3-demo-bucket/myhudidataset/')
```

> **Note**
>
> Você pode ver “hoodie” em vez de Hudi em exemplos de código e notificações. A base de código do Hudi usa amplamente a antiga grafia “hoodie”.

DataSourceWriteOptions referência para o Hudi

Opção	Descrição
TABLE_NAME	O nome da tabela com o qual registrar o conjunto de dados.
TABLE_TYPE_OPT_KEY	Opcional. Especifica se o conjunto de dados foi criado como `"COPY_ON_WRITE"` ou `"MERGE_ON_READ"` . O padrão é `"COPY_ON_WRITE"` .
RECORDKEY_FIELD_OPT_KEY	O campo de chave de registro cujo valor será usado como o componente `recordKey` de `HoodieKey` . O valor real será obtido invocando `.toString()` no valor do campo. Campos aninhados podem ser especificados usando a notação de pontos, por exemplo, `a.b.c`.
PARTITIONPATH_FIELD_OPT_KEY	O campo de caminho de partição cujo valor será usado como o componente `partitionPath` de `HoodieKey` . O valor real será obtido invocando `.toString()` no valor do campo.
PRECOMBINE_FIELD_OPT_KEY	O campo usado na pré-combinação antes da gravação real. Quando dois registros têm o mesmo valor de chave, o Hudi seleciona aquele com o maior valor para o campo de pré-combinação, conforme determinado por `Object.compareTo(..)` .

As opções a seguir são necessárias apenas para registrar a tabela do conjunto de dados do Hudi no seu metastore. Se você não registrar o conjunto de dados do Hudi como uma tabela no metastore do Hive, essas opções não serão necessárias.

DataSourceWriteOptions referência para o Hive

Opção	Descrição
HIVE_DATABASE_OPT_KEY	O banco de dados do Hive com o qual sincronizar. O padrão é `"default"` .
HIVE_PARTITION_EXTRACTOR_CLASS_OPT_KEY	A classe usada para extrair valores de campo de partição em colunas de partição do Hive.
HIVE_PARTITION_FIELDS_OPT_KEY	O campo no conjunto de dados a ser usado para determinar colunas de partição do Hive.
HIVE_SYNC_ENABLED_OPT_KEY	Quando definido como `"true"`, registra o conjunto de dados no metastore do Apache Hive. O padrão é `"false"`.
HIVE_TABLE_OPT_KEY	Obrigatório. O nome da tabela no Hive com a qual sincronizar. Por exemplo, .`"my_hudi_table_cow"`
HIVE_USER_OPT_KEY	Opcional. O nome de usuário do Hive a ser usado ao sincronizar. Por exemplo, .`"hadoop"`
HIVE_PASS_OPT_KEY	Opcional. A senha do Hive para o usuário especificado por `HIVE_USER_OPT_KEY` .
HIVE_URL_OPT_KEY	O URL do metastore do Hive.

Upsert dados

O exemplo a seguir demonstra como inserir novos dados gravando um. DataFrame Ao contrário do exemplo anterior de inserção, o valor `OPERATION_OPT_KEY` é definido como `UPSERT_OPERATION_OPT_VAL`. Além disso, `.mode(SaveMode.Append)` é especificado para indicar que o registro deve ser anexado.

Upsert usando Scala com as versões 6.7 e posteriores do Amazon EMR

> **Note**
>
> O Amazon EMR 6.7.0 usa o Apache Hudi 0.11.0-amzn-0, que contém melhorias significativas em relação às versões anteriores do Hudi. Para obter mais informações, consulte o Guia de migração do Apache Hudi 0.11.0. Os exemplos nesta guia refletem essas alterações.

```
// Create a new DataFrame from the first row of inputDF with a different creation_date
 value
val updateDF = inputDF.limit(1).withColumn("creation_date", lit("new_value"))

(updateDF.write
    .format("hudi")
    .options(hudiOptions)
    .option(DataSourceWriteOptions.OPERATION_OPT_KEY, "upsert")
    .mode(SaveMode.Append)
    .save("s3://amzn-s3-demo-bucket/myhudidataset/"))
```

Upsert usando Scala com as versões 6.6 e anteriores do Amazon EMR

```
// Create a new DataFrame from the first row of inputDF with a different creation_date
 value
val updateDF = inputDF.limit(1).withColumn("creation_date", lit("new_value"))

(updateDF.write
    .format("org.apache.hudi")
    .option(DataSourceWriteOptions.OPERATION_OPT_KEY,
 DataSourceWriteOptions.UPSERT_OPERATION_OPT_VAL)
    .options(hudiOptions)
    .mode(SaveMode.Append)
    .save("s3://amzn-s3-demo-bucket/myhudidataset/"))
```

Upsert usando PySpark

```
from pyspark.sql.functions import lit

# Create a new DataFrame from the first row of inputDF with a different creation_date
 value
```

```
updateDF = inputDF.limit(1).withColumn('creation_date', lit('new_value'))

updateDF.write \
    .format('org.apache.hudi') \
    .option('hoodie.datasource.write.operation', 'upsert') \
    .options(**hudiOptions) \
    .mode('append') \
    .save('s3://amzn-s3-demo-bucket/myhudidataset/')
```

Excluir um registro

Para excluir um registro de forma irreversível, você pode upsert uma carga útil vazia. Nesse caso, a opção `PAYLOAD_CLASS_OPT_KEY` especifica a classe `EmptyHoodieRecordPayload`. O exemplo usa o mesmo DataFrame,`updateDF`, usado no exemplo de upsert para especificar o mesmo registro.

Excluir usando Scala com as versões 6.7 e posteriores do Amazon EMR

> **Note**
>
> O Amazon EMR 6.7.0 usa o Apache Hudi 0.11.0-amzn-0, que contém melhorias significativas em relação às versões anteriores do Hudi. Para obter mais informações, consulte o Guia de migração do Apache Hudi 0.11.0. Os exemplos nesta guia refletem essas alterações.

```
(updateDF.write
    .format("hudi")
    .options(hudiOptions)
    .option(DataSourceWriteOptions.OPERATION_OPT_KEY, "delete")
    .mode(SaveMode.Append)
    .save("s3://amzn-s3-demo-bucket/myhudidataset/"))
```

Excluir usando Scala com as versões 6.6 e anteriores do Amazon EMR

```
(updateDF.write
    .format("org.apache.hudi")
    .option(DataSourceWriteOptions.OPERATION_OPT_KEY,
 DataSourceWriteOptions.UPSERT_OPERATION_OPT_VAL)
    .option(DataSourceWriteOptions.PAYLOAD_CLASS_OPT_KEY,
 "org.apache.hudi.common.model.EmptyHoodieRecordPayload")
```

```
    .mode(SaveMode.Append)
    .save("s3://amzn-s3-demo-bucket/myhudidataset/"))
```

Excluir usando PySpark

```
updateDF.write \
    .format('org.apache.hudi') \
    .option('hoodie.datasource.write.operation', 'upsert') \
    .option('hoodie.datasource.write.payload.class',
 'org.apache.hudi.common.model.EmptyHoodieRecordPayload') \
    .options(**hudiOptions) \
    .mode('append') \
    .save('s3://amzn-s3-demo-bucket/myhudidataset/')
```

Você também pode excluir dados de forma irreversível definindo `OPERATION_OPT_KEY` como `DELETE_OPERATION_OPT_VAL` para remover todos os registros no conjunto de dados enviado. Para obter instruções sobre como realizar exclusões reversíveis e obter mais informações sobre a exclusão de dados armazenados em tabelas do Hudi, consulte Exclusões na documentação do Apache Hudi.

Ler em um conjunto de dados do Hudi

Para recuperar dados no momento atual, o Hudi realiza consultas de snapshots por padrão. Veja a seguir um exemplo de consulta do conjunto de dados gravado no S3 em Gravar em um conjunto de dados do Hudi. *s3://amzn-s3-demo-bucket/myhudidataset*Substitua pelo caminho da tabela e adicione asteriscos curinga para cada nível de partição, além de um asterisco adicional. Neste exemplo, há um nível de partição, portanto adicionamos dois símbolos curinga.

Ler usando Scala com as versões 6.7 e posteriores do Amazon EMR

> **Note**
>
> O Amazon EMR 6.7.0 usa o Apache Hudi 0.11.0-amzn-0, que contém melhorias significativas em relação às versões anteriores do Hudi. Para obter mais informações, consulte o Guia de migração do Apache Hudi 0.11.0. Os exemplos nesta guia refletem essas alterações.

```
val snapshotQueryDF = spark.read
    .format("hudi")
    .load("s3://amzn-s3-demo-bucket/myhudidataset")
```

```
    .show()
```

Ler usando Scala com as versões 6.6 e anteriores do Amazon EMR

```
(val snapshotQueryDF = spark.read
    .format("org.apache.hudi")
    .load("s3://amzn-s3-demo-bucket/myhudidataset" + "/*/*"))

snapshotQueryDF.show()
```

Leia usando PySpark

```
snapshotQueryDF = spark.read \
    .format('org.apache.hudi') \
    .load('s3://amzn-s3-demo-bucket/myhudidataset' + '/*/*')

snapshotQueryDF.show()
```

Consultas incrementais

Você também pode realizar consultas incrementais com o Hudi para obter um fluxo de registros que foram alterados desde um determinado carimbo de data/hora de confirmação. Para fazer isso, defina o campo `QUERY_TYPE_OPT_KEY` como `QUERY_TYPE_INCREMENTAL_OPT_VAL`. Em seguida, adicione um valor para `BEGIN_INSTANTTIME_OPT_KEY` para obter todos os registros gravados desde a hora especificada. Normalmente, as consultas incrementais são dez vezes mais eficientes do que as de lote, pois processam somente registros alterados.

Ao realizar consultas incrementais, use o caminho da tabela raiz (básica) sem os asteriscos curinga usados nas consultas Snapshot.

> **Note**
>
> O Presto não é compatível com consultas incrementais.

Consultas incrementais usando o Scala

```
(val incQueryDF = spark.read
    .format("org.apache.hudi")
    .option(DataSourceReadOptions.QUERY_TYPE_OPT_KEY,
 DataSourceReadOptions.QUERY_TYPE_INCREMENTAL_OPT_VAL)
```

```
    .option(DataSourceReadOptions.BEGIN_INSTANTTIME_OPT_KEY, <beginInstantTime>)
    .load("s3://amzn-s3-demo-bucket/myhudidataset" ))

incQueryDF.show()
```

Consultas incrementais usando PySpark

```
readOptions = {
  'hoodie.datasource.query.type': 'incremental',
  'hoodie.datasource.read.begin.instanttime': <beginInstantTime>,
}

incQueryDF = spark.read \
    .format('org.apache.hudi') \
    .options(**readOptions) \
    .load('s3://amzn-s3-demo-bucket/myhudidataset')

incQueryDF.show()
```

Para obter mais informações sobre a leitura de conjuntos de dados do Hudi, consulte Consultar tabelas do Hudi na documentação do Apache Hudi.

Usar a CLI do Hudi

É possível usar a CLI do Hudi para administrar conjuntos de dados do Hudi para visuaalizar informações sobre confirmações, o sistema de arquivos, estatísticas e muito mais. Você também pode usar a CLI para executar manualmente compactações, programá-las ou cancelar as compactações programadas. Para obter mais informações, consulte Interacting via CLI na documentação do Apache Hudi.

Iniciar a CLI do Hudi e conectar-se a um conjunto de dados

1. Conecte-se ao nó principal usando SSH. Para obter mais informações, consulte Conectar-se ao nó principal usando SSH no Guia de gerenciamento do Amazon EMR.
2. Na linha de comando, digite `/usr/lib/hudi/cli/bin/hudi-cli.sh`.

   O prompt de comando muda para `hudi->`.
3. Use o comando a seguir para se conectar a um conjunto de dados. *s3://amzn-s3-demo-bucket/myhudidataset*Substitua pelo caminho para o conjunto de dados com o qual você deseja trabalhar. O valor que usamos é o mesmo estabelecido em exemplos anteriores.

```
connect --path s3://amzn-s3-demo-bucket/myhudidataset
```

O prompt de comando é alterado para incluir o conjunto de dados ao qual você está conectado, conforme mostrado no exemplo a seguir.

```
hudi:myhudidataset->
```

Por padrão, o `hudi-cli.sh` script na versão 7.3.0 a 7.8.0 do Amazon EMR usa. `hudi-cli-bundle.jar` Caso tenha problemas, é possível voltar para a CLI clássica do Hudi com o seguinte comando:

```
/usr/lib/hudi/cli/bin/hudi-cli.sh --cliBundle false
```

Esse comando executa o script `hudi-cli.sh`, define o sinalizador `--cliBundle` e instrui a CLI a usar os arquivos JAR individuais em vez do JAR empacotado. Por padrão, o `--cliBundle` é definido como verdadeiro, o que significa que a CLI usa o JAR empacotado.

Usando o Amazon EMR 7.9.0 e versões posteriores

> **Note**
>
> O script hudi-cli.sh foi descontinuado na versão 7.9.0 e superior do EMR. A versão 7.9.0 e superior do Amazon EMR usa .jar. hudi-cli-bundle

Iniciar a CLI do Hudi e conectar-se a um conjunto de dados:

1. Conecte-se ao nó principal usando SSH. Para obter mais informações, consulte Conectar-se ao nó principal usando SSH no Guia de gerenciamento do Amazon EMR.
2. Na linha de comando, digite/usr/lib/hudi/cli-bundle/bin/hudi- cli-with-bundle .sh ou simplesmente digite hudi-cli-with-bundleou >hudi-cli.

   O prompt de comando muda para hudi- >.
3. Use o comando a seguir para se conectar a um conjunto de dados. Substitua s3://amzn-s3-demo-bucket/myhudidataset pelo caminho para o conjunto de dados com o qual você deseja trabalhar. O valor que usamos é o mesmo estabelecido em exemplos anteriores.

```
connect --path s3://amzn-s3-demo-bucket/myhudidataset
```

4. O prompt de comando é alterado para incluir o conjunto de dados ao qual você está conectado, conforme mostrado no exemplo a seguir.

```
hudi:myhudidataset->
```

Histórico de versões do Hudi

A tabela a seguir informa a versão do Hudi incluída em cada versão do Amazon EMR, além dos componentes instalados com a aplicação. Para obter as versões dos componentes em cada versão, consulte a seção Versão de componentes da sua versão em Versões 7.x do Amazon EMR, Versões de lançamento 6.x do Amazon EMR ou Versões de lançamento 5.x do Amazon EMR.

Informações de versão do Hudi

Rótulo de versão do Amazon EMR	Versão do Hudi	Componentes instalados com o Hudi
emr-7.9.0	0.15.0-amzn-6	Not available.
emr-7.8.0	0.15.0-amzn-5	Not available.
emr-7.7.0	0.15.0-amzn-4	Not available.
emr-7.6.0	0.15.0-amzn-3	Not available.
emr-7.5.0	0.15.0-amzn-2	Not available.
emr-7.4.0	0.15.0-amzn-1	Not available.
emr-7.3.0	0.15.0-amzn-0	Not available.
emr-7.2.0	0.14.1-amzn-1	Not available.
emr-5.36.2	0.10.1-amzn-1	Not available.
emr-7.1.0	0.14.1-amzn-0	Not available.

Rótulo de versão do Amazon EMR	Versão do Hudi	Componentes instalados com o Hudi
emr-7.0.0	0.14.0-amzn-1	Not available.
emr-6.15.0	0.14.0-amzn-0	Not available.
emr-6.14.0	0.13.1-amzn-2	Not available.
emr-6.13.0	0.13.1-amzn-1	Not available.
emr-6.12.0	0.13.1-amzn-0	Not available.
emr-6.11.1	0.13.0-amzn-0	Not available.
emr-6.11.0	0.13.0-amzn-0	Not available.
emr-6.10.1	0.12.2-amzn-0	Not available.
emr-6.10.0	0.12.2-amzn-0	Not available.
emr-6.9.1	0.12.1-amzn-0	Not available.
emr-6.9.0	0.12.1-amzn-0	Not available.
emr-6.8.1	0.11.1-amzn-0	Not available.
emr-6.8.0	0.11.1-amzn-0	Not available.
emr-6.7.0	0.11.0-amzn-0	Not available.
emr-5.36.1	0.10.1-amzn-1	Not available.
emr-5.36.0	0.10.1-amzn-1	Not available.
emr-6.6.0	0.10.1-amzn-0	Not available.
emr-5.35.0	0.9.0-amzn-2	Not available.
emr-6.5.0	0.9.0-amzn-1	Not available.
emr-6.4.0	0.8.0-amzn-0	Not available.

Rótulo de versão do Amazon EMR	Versão do Hudi	Componentes instalados com o Hudi
emr-6.3.1	0.7.0-amzn-0	Not available.
emr-6.3.0	0.7.0-amzn-0	Not available.
emr-6.2.1	0.6.0-amzn-1	Not available.
emr-6.2.0	0.6.0-amzn-1	Not available.
emr-6.1.1	0.5.2-incubating-amzn-2	Not available.
emr-6.1.0	0.5.2-incubating-amzn-2	Not available.
emr-6.0.1	0.5.0-incubating-amzn-1	Not available.
emr-6.0.0	0.5.0-incubating-amzn-1	Not available.
emr-5.34.0	0.9.0-amzn-0	Not available.
emr-5.33.1	0.7.0-amzn-1	Not available.
emr-5.33.0	0.7.0-amzn-1	Not available.
emr-5.32.1	0.6.0-amzn-0	Not available.
emr-5.32.0	0.6.0-amzn-0	Not available.
emr-5.31.1	0.6.0-amzn-0	Not available.
emr-5.31.0	0.6.0-amzn-0	Not available.
emr-5.30.2	0.5.2-incubating	Not available.
emr-5.30.1	0.5.2-incubating	Not available.
emr-5.30.0	0.5.2-incubating	Not available.
emr-5.29.0	0.5.0-incubating	Not available.
emr-5.28.1	0.5.0-incubating	Not available.

Rótulo de versão do Amazon EMR	Versão do Hudi	Componentes instalados com o Hudi
emr-5.28.0	0.5.0-incubating	Not available.

Hue

O Hue (Hadoop User Experience) é uma interface gráfica do usuário com base na Web de código-fonte aberto para uso com o Amazon EMR e o Apache Hadoop. O Hue agrupa vários projetos de ecossistemas do Hadoop diferentes em uma interface configurável. O Amazon EMR também adicionou personalizações específicas para o Hue no Amazon EMR. O Hue atua como um front-end para aplicativos executados no cluster, permitindo que você interaja com os aplicativos usando uma interface que possa ser mais familiar ou amigável para o usuário. O aplicativos no Hue, como os editores do Hive e do Pig, dispensam a necessidade de fazer login no cluster para executar scripts interativamente usando o shell de cada aplicativo. Depois que um cluster for iniciado, você pode interagir totalmente com os aplicativos usando o Hue ou uma interface similar. Para obter mais informações sobre o Hue, consulte http://gethue.com.

Por padrão, o Hue é instalado quando você inicia um cluster usando o console do Amazon EMR. Você pode optar por não instalar o Hue ao usar Opções avançadas no console do Amazon EMR quando iniciar um cluster ou ao especificar explicitamente a opção `--applications` e omitindo o Hue, quando usar `create-cluster` na AWS CLI.

Tópicos

- Informações de versão do Hue
- Atributos compatíveis e incompatíveis do Hue no Amazon EMR
- Considerações
- Conectar-se à interface de usuário da Web do Hue
- Usar o Hue com um banco de dados remoto no Amazon RDS
- Configurações avançadas do Hue
- Histórico de versões do Hue

Informações de versão do Hue

Versão do Hue para 7.9.0

A tabela a seguir lista a versão do Hue incluída na versão mais recente da série 7.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Hue.

Para obter a versão dos componentes instalados com o Hue nesse lançamento, consulte Versões dos componentes da versão 7.9.0.

Informações de versão do Hue para o emr-7.9.0

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-7.9.0	Hue 4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Versão do Hue para 6.15.0

A tabela a seguir lista a versão do Hue incluída na versão mais recente da série 6.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Hue.

Para obter a versão dos componentes instalados com o Hue nessa versão, consulte Release 6.15.0 Component Versions.

Informações de versão do Hue para o emr-6.15.0

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-6.15.0	Hue 4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
		timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Versão do Hue para 5.36.2

A tabela a seguir lista a versão do Hue incluída na versão mais recente da série 5.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Hue.

Para obter a versão dos componentes instalados com o Hue nesse lançamento, consulte Release 5.36.2 Component Versions.

Informações de versão do Hue para o emr-5.36.2

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.36.2	Hue 4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Atributos compatíveis e incompatíveis do Hue no Amazon EMR

Com o Amazon EMR 7.0.0 e versões superiores, o Hue exige o Python 3.9 ou superior. Ao usar criptografia em trânsito, é necessário usar uma chave de cifra de 2048 bits.

- Amazon S3 e Hadoop File System (HDFS) Browser

  - Com as permissões apropriadas, você pode navegar e mover dados entre o armazenamento HDFS temporário e buckets do S3 pertencentes à sua conta.
  - Por padrão, os superusuários do Hue podem acessar todos os arquivos que os perfis do IAM do Amazon EMR têm permissão para acessar. Os usuários recém-criados não têm automaticamente permissões para acessar o navegador de arquivos Amazon S3 e devem ter as permissões `filebrowser.s3_access` ativadas para o grupo deles.
- Hive: execute consultas interativas nos seus dados. Esta também é uma maneira útil de gerar protótipos de consultas programáticas ou em lotes.
- Pig: execute scripts nos seus dados ou emita comandos interativos.
- Oozie: crie e monitore fluxos de trabalho do Oozie.
- Metastore Manager: visualize e manipule o conteúdo do Hive Metastore (importar,criar, descartar e assim por diante).
- Navegador de trabalhos: veja o status dos seus trabalhos do Hadoop enviados.
- Gerenciamento de usuários: gerencie contas do Hue e integre usuários do LDAP ao Hue.
- AWS Exemplos da: há vários ready-to-run exemplos de dados de vários AWS serviços da usando aplicações no Hue. Ao fazer logon no Hue, você será levado ao aplicativo Hue Home, onde os exemplos estão pré-instalados.
- O servidor Livy só é compatível com as versões 5.9.0 e posteriores do Amazon EMR.
- Para usar o Hue Notebook para Spark, você deve instalar o Hue com Livy e Spark.
- Trino/Presto: o Hue oferece suporte à conexão com os servidores de dados Trino e Presto. Com o Amazon EMR 7.0.0 e versões superiores, essa conectividade é facilitada com os conectores `trino-python-client` e `presto-python-client`. Versões inferiores à 7.0.0 usam um conector JDBC por meio de um gateway JDBC.
- Não há suporte para o painel do Hue.
- O PostgreSQL não é compatível.

Considerações

Considere as limitações a seguir ao usar o Hue no Amazon EMR.

Performance com grandes tabelas de metadados do Hue

Se o banco de dados de metadados do Hue ficar muito grande, a performance poderá diminuir. Para verificar o tamanho das tabelas, primeiro conecte-se ao nó primário do EC2 cluster do Amazon

EMR no Amazon com SSH e execute o comando `sudo mysql -u root` para iniciar a CLI do MySQL. Para obter o tamanho da sua tabela, execute a consulta `SELECT COUNT(*) FROM hue.<table_name>`. Confira o que `<table_name>` pode ser:

- desktop_document
- desktop_document2
- oozie_job
- beeswax_savedquery
- beeswax_session
- beeswax_queryhistory

Se a execução dessa consulta retornar uma contagem superior a 100.000, você deverá executar o comando de limpeza a seguir para excluir os registros antigos.

```
cd /opt/cloudera/parcels/CDH/lib/hue # Hue home directory ./build/env/bin/hue
 desktop_document_cleanup
```

Para obter mais informações sobre como limpar seu banco de dados, consulte a página de referência na documentação do Hue.

O Hue não limpa automaticamente as tabelas, mas as versões 5.12.0 e posteriores do Amazon EMR fornecem um método para excluir documentos antigos nelas. Crie o script de shell a seguir e execute-o como uma etapa em um cluster do Amazon EMR com um parâmetro inteiro que representa o número máximo de dias para manter documentos no banco de dados.

```
#!/bin/bash
if grep isMaster /mnt/var/lib/info/instance.json | grep false;
then
  echo "This is not the primary node; do nothing, exiting"
  exit 0
fi
while [ ! -f /usr/lib/hue/desktop/core/src/desktop/management/commands/
desktop_document_cleanup.py ]
do
  sleep 1
done
sudo systemctl stop hue.service
sudo sed -i 's+  LOG.warn+  # LOG.warn+g' /usr/lib/hue/desktop/core/src/desktop/
management/commands/desktop_document_cleanup.py
```

```
sudo /usr/lib/hue/build/env/bin/hue desktop_document_cleanup --keep-days $1
sudo systemctl start hue.service
```

Incompatibilidade entre versões do Hue

Caso esteja usando o mesmo banco de dados de metadados do Hue em vários clusters habilitados pelo Hue, recomendamos que esses clusters executem a mesma versão do Hue. Versões diferentes do Hue podem ter esquemas distintos para o banco de dados de metadados do Hue. Usar o mesmo banco de dados para versões diferentes pode causar falha na instalação do Hue. Por exemplo, tentar usar o mesmo banco de dados para dois clusters com as versões 4.10.0 e 4.11.0 instaladas pode causar erros de login para os usuários que estão tentando entrar no cluster 4.10.0.

Para obter uma lista dos rótulos de lançamento do Amazon EMR e das versões instaladas correspondentes do Hue, consulte o histórico de lançamentos do Hue.

Conectar-se à interface de usuário da Web do Hue

Conectar-se à interface de usuário da Web do Hue equivale a conectar-se a qualquer interface HTTP hospedada no nó principal de um cluster. O procedimento a seguir descreve como acessar a interface de usuário do Hue. Para obter mais informações, consulte Visualizar interfaces da Web hospedadas em clusters do EMR no Guia de gerenciamento do Amazon EMR.

Visualizar a interface de usuário da Web do Hue

1. Siga estas instruções para Configurar um túnel SSH para o nó principal usando o encaminhamento de portas dinâmicas no Guia de gerenciamento do Amazon EMR.
2. Digite o seguinte endereço no seu navegador para abrir a interface web do Hue: `http://`*`master public DNS`*`:8888` onde *`master public dns`* está o nome DNS público do seu nó principal do cluster, por exemplo. `ec2-11-22-333-44.compute-1.amazonaws.com`
3. Na tela de login do Hue, se você for o administrador fazendo login pela primeira vez, insira um nome de usuário e uma senha para criar sua superconta no Hue e selecione Criar conta. Caso contrário, digite seu nome de usuário e sua senha e selecione Create account (Criar conta) ou insira as credenciais fornecidas pelo administrador.

Usar o Hue com um banco de dados remoto no Amazon RDS

Por padrão, as informações de usuários do Hue e os histórico de consultas são armazenados em um banco de dados MySQL local no nó principal. Como opção, você pode criar um ou mais clusters habilitados para o Hue usando uma configuração armazenada no Amazon S3 e um banco de dados MySQL no Amazon Relational Database Service (Amazon RDS). Isso permite que você mantenha as informações dos usuários e o histórico de consultas criados pelo Hue sem manter o cluster do Amazon EMR em execução. Recomendamos o uso da criptografia do lado do servidor do Amazon S3 para armazenar o arquivo de configuração.

Primeiro, crie o banco de dados remoto para o Hue.

Para criar o banco de dados MySQL externo

1. Abra o console do Amazon RDS em https://console.aws.amazon.com/rds/.
2. Clique em Launch a DB Instance (Iniciar uma instância de banco de dados).
3. Escolha MySQL e clique em Select (Selecionar).
4. Deixe a seleção padrão de Multi-AZ Deployment and Provisioned IOPS Storage (Implantação Multi-AZ e armazenamento de IOPS provisionadas) e clique em Next (Próximo).
5. Deixe as especificações da instância em seus valores padrão, especifique configurações e clique em Next (Próximo).
6. Na página Configure Advanced Settings, escolha um nome de security group apropriado e o nome do banco de dados. O security group que você usar deve permitir pelo menos o acesso TCP de entrada para a porta 3306 a partir do nó principal do seu cluster. Se você não tiver criado o cluster neste ponto, poderá permitir que todos os hosts se conectem à porta 3306 e ajustar o security group depois de ter executado o cluster. Clique em Launch DB Instance (Iniciar instância de banco de dados).
7. Em RDS Dashboard (Painel do RDS), selecione Instances (Instâncias) e selecione a instância que você acabou de criar. Quando seu banco de dados estiver disponível, anote o nome do banco de dados, o nome de usuário, a senha e o nome de host da instância do RDS. Você usará essas informações ao criar e configurar o cluster.

Para especificar um banco de dados MySQL externo para o Hue ao iniciar um cluster usando o AWS CLI

Para especificar um banco de dados MySQL externo para o Hue ao iniciar um cluster usando o AWS CLI, use as informações que você observou ao criar sua instância do RDS para configuração com um objeto de configuraçãohue.ini.

> **Note**
>
> Você pode criar vários clusters que usam o mesmo banco de dados externo, mas cada cluster compartilhará informações de usuários e o histórico de usuários.

- Usando o AWS CLI, crie um cluster com o Hue instalado, usando o banco de dados externo que você criou e referenciando um arquivo de configuração com uma classificação de configuração para o Hue que especifica as propriedades do banco de dados. O exemplo a seguir cria um cluster com o Hue instalado, fazendo referência a um arquivo de configuração no Amazon S3, myConfig.json, que especifica a configuração do banco de dados.

  > **Note**
  >
  > Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

  ```
  aws emr create-cluster --release-label emr-7.9.0 --applications Name=Hue Name=Spark
   Name=Hive \
  --instance-type m5.xlarge --instance-count 3 \
  --configurations https://s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/
  myConfig.json --use-default-roles
  ```

  Os conteúdos de exemplo do arquivo myConfig.json são mostrados abaixo. Substitua *dbnameusername*,*password*,, e *RDS instance hostname* pelos valores que você anotou anteriormente no Painel do RDS.

  ```
  [{
    "Classification": "hue-ini",
    "Properties": {},
  ```

```
  "Configurations": [
    {
      "Classification": "desktop",
      "Properties": {},
      "Configurations": [
        {
          "Classification": "database",
          "Properties": {
            "name": "dbname",
            "user": "username",
            "password": "password",
            "host": "RDS instance hostname",
            "port": "3306",
            "engine": "mysql"
          },
          "Configurations": []
        }
      ]
    }
  ]
}]
```

Solução de problemas

No caso de um failover do Amazon RDS

É possível que os usuários encontrem atrasos ao executar uma consulta, pois a instância de banco de dados do Hue não está respondendo ou está em processo de failover. Veja a seguir estão alguns fatos e diretrizes para esse problema:

- Se você fizer login no console do Amazon RDS, poderá procurar eventos de failover. Por exemplo, para ver se um failover é em andamento ou já ocorreu, procure eventos como "Multi-AZ instance failover started" e "Multi-AZ instance failover completed".
- Leva cerca de 30 segundos para uma instância do RDS concluir um failover.
- Se você estiver recebendo longer-than-normal respostas para consultas no Hue, tente voltar a executar a consulta.

Configurações avançadas do Hue

Esta seção inclui os seguintes tópicos.

Tópicos

- Configurar o Hue para usuários do LDAP

Configurar o Hue para usuários do LDAP

A integração com o LDAP permite que os usuários façam login no Hue usando credenciais existentes armazenadas em um diretório LDAP. Ao integrar o Hue com o LDAP, você não precisa gerenciar independentemente as informações dos usuários no Hue. As informações abaixo demonstram a integração do Hue com o Microsoft Active Directory, mas as opções de configuração são análogas a qualquer diretório LDAP.

A autenticação LDAP primeiro vincula-se ao servidor e estabelece a conexão. Em seguida, a conexão estabelecida é usada para consultas subsequentes, para procurar informações de usuários LDAP. A menos que o seu servidor do Active Directory permita conexões anônimas, uma conexão precisa ser estabelecida usando um nome diferenciado e uma senha de associação. O nome diferenciado (ou DN) de associação é definido pela configuração `bind_dn`. A senha de associação é definida pela configuração `bind_password`. O Hue tem duas maneiras de associar solicitações LDAP: associação de pesquisa e associação direta. O método preferencial para o uso do Hue com o Amazon EMR é a associação de pesquisa.

Quando a associação de pesquisa é usada com o Active Directory, o Hue usa o atributo de nome de usuário (definido por `user_name_attr config`) para localizar o atributo que precisa ser recuperado do nome diferenciado (ou DN) base. A associação de pesquisa é útil quando o DN completo não é conhecido pelo usuário do Hue.

Por exemplo, você pode ter `user_name_attr config` definido para usar o nome comum (ou CN). Nesse caso, o servidor do Active Directory usa o nome de usuário do Hue fornecido durante o login para pesquisar a árvore de diretório em busca de um nome comum correspondente, começando com o nome diferenciado base. Se o nome comum do usuário do Hue for encontrado, o nome diferenciado desse usuário será retornado pelo servidor. Em seguida, o Hue constrói um nome diferenciado usado para autenticar o usuário, realizando uma operação de associação.

> **Note**
>
> A associação de pesquisa procura nomes de usuários em todas as subárvores de diretórios, começando com o nome diferenciado base. O nome diferenciado base especificado na configuração LDAP do Hue deve ser o pai mais próximo do nome do usuário, ou o desempenho da autenticação LDAP poderá ser prejudicado.

Quando a associação direta é usada com o Active Directory, o `nt_domain` ou `ldap_username_pattern` exato deve ser usado para se autenticar. Quando a associação direta é usada, se o atributo de domínio nt (definido pela configuração `nt_domain`) estiver especificado, um template de nome diferenciado do usuário será criado usando o formato: `<login username>@nt_domain`. Esse template é usado para pesquisar todas as subárvores de diretórios, começando com o nome diferenciado base. Se o domínio nt não for configurado, o Hue procurará um padrão de nome diferenciado exato para o usuário (definido pela configuração `ldap_username_pattern`). Nesse caso, o servidor procura um valor `ldap_username_pattern` correspondente em todas as subárvores de diretórios, começando com o nome diferenciado base.

Iniciar um cluster com propriedades do LDAP para o Hue usando a AWS CLI

- Para especificar propriedades do LDAP para `hue-ini`, crie um cluster com o Hue instalado e faça referência a um arquivo json com propriedades de configuração para o LDAP. Um exemplo de comando é mostrado abaixo, que faz referência a um arquivo de configuração `myConfig.json` armazenado no Amazon S3.

  ```
  aws emr create-cluster --release-label emr-7.9.0 --applications Name=Hue Name=Spark
   Name=Hive \
  --instance-type m5.xlarge --instance-count 3 --configurations https://
  s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myConfig.json.
  ```

  Os conteúdos de exemplo de `myConfig.json` são mostrados abaixo.

  ```
  [
      {
          "Classification": "hue-ini",
          "Properties": {},
          "Configurations": [
              {
                  "Classification": "desktop",
  ```

```
                "Properties": {},
                "Configurations": [
                    {
                        "Classification": "ldap",
                        "Properties": {},
                        "Configurations": [
                            {
                                "Classification": "ldap_servers",
                                "Properties": {},
                                "Configurations": [
                                    {
                                        "Classification": "yourcompany",
                                        "Properties": {
                                            "base_dn":
 "DC=yourcompany,DC=hue,DC=com",
                                            "ldap_url": "ldap://ldapurl",
                                            "search_bind_authentication": "true",
                                            "bind_dn":
 "CN=hue,CN=users,DC=yourcompany,DC=hue,DC=com",
                                            "bind_password": "password"
                                        },
                                        "Configurations": []
                                    }
                                ]
                            }
                        ]
                    },
                    {
                        "Classification": "auth",
                        "Properties": {
                         "backend": "desktop.auth.backend.LdapBackend"
                              }
                    }
                ]
            }

        ]
    }
]
```

> **Note**
>
> Com as versões 5.21.0 e posteriores do Amazon EMR, você pode substituir as configurações de cluster e especificar classificações de configuração adicionais para cada grupo de instâncias em um cluster em execução. Você faz isso usando o console do Amazon EMR, o AWS Command Line Interface (AWS CLI) ou o AWS SDK. Para obter mais informações, consulte Supplying a Configuration for an Instance Group in a Running Cluster.

Visualizar as configurações do LDAP no Hue

1. Verifique se você tem uma conexão VPN ou túnel SSH ativo para o nó principal do cluster do Amazon EMR. Em seguida, no seu navegador, digite:8888 *master-public-dns* para abrir a interface web do Hue.
2. Faça login usando suas credenciais de administrador do Hue. Se a janela Did you know? (Você sabia?) for aberta, clique em Got it, prof! (Entendido!) para fechá-la.
3. Clique no ícone do Hue na barra de ferramentas.
4. Na página Sobre o Hue, selecione Configuração.
5. Na seção Configuration Sections and Variables (Seções e variáveis de configuração), clique em Desktop.
6. Role até a seção ldap para exibir as configurações.

Histórico de versões do Hue

A tabela a seguir lista a versão do Hue incluída em cada versão do Amazon EMR, além dos componentes instalados com a aplicação. Para obter as versões dos componentes em cada versão, consulte a seção Versão de componentes da sua versão em Versões 7.x do Amazon EMR, Versões de lançamento 6.x do Amazon EMR ou Versões de lançamento 5.x do Amazon EMR.

Informações de versão do Hue

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-7.9.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode,

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
		hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-7.8.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-7.7.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-7.6.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-7.5.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-7.4.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-7.3.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-7.2.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.36.2	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-7.1.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-7.0.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-6.15.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.14.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-6.13.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.12.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-6.11.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.11.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-6.10.1	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.10.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-6.9.1	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.9.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-6.8.1	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.8.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-6.7.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.36.1	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.36.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.6.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.35.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.5.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-6.4.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.3.1	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-6.3.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.2.1	4.8.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-6.2.0	4.8.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.1.1	4.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-6.1.0	4.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.0.1	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-6.0.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.34.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.33.1	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.33.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.32.1	4.8.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.32.0	4.8.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.31.1	4.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.31.0	4.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.30.2	4.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.30.1	4.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.30.0	4.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.29.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.28.1	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.28.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.27.1	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.27.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.26.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.25.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.24.1	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.24.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.23.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.23.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.22.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.21.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.21.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.21.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.20.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.20.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.19.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.19.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.18.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.18.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.17.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.17.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.17.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.16.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.16.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.15.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.15.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.14.2	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.14.1	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.14.0	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.13.1	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.13.0	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.12.3	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.12.2	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.12.1	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.12.0	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.11.4	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.11.3	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.11.2	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.11.1	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.11.0	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.10.1	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.10.0	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.9.1	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.9.0	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.8.3	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.8.2	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.8.1	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.8.0	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.7.1	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.7.0	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.6.1	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.6.0	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.5.4	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.5.3	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.5.2	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.5.1	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.5.0	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.4.1	3.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.4.0	3.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.3.2	3.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.3.1	3.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.3.0	3.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.2.3	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.2.2	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.2.1	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.2.0	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.1.1	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.1.0	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.0.3	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-5.0.2	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-5.0.1	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-5.0.0	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-4.9.6	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.9.5	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.9.4	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-4.9.3	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.9.2	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.9.1	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-4.8.5	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.8.4	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.8.3	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-4.8.2	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.8.1	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.8.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-4.7.4	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.7.3	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.7.2	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-4.7.1	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.7.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.6.1	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-4.6.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.5.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.4.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server

Rótulo de versão do Amazon EMR	Versão do Hue	Componentes instalados com o Hue
emr-4.3.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.2.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.1.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server

Iceberg

O Apache Iceberg é um formato de tabela aberta para grandes conjuntos de dados no Amazon Simple Storage Service (Amazon S3). Ele fornece performance rápida de consultas em tabelas grandes, confirmações atômicas, gravações simultâneas e evolução de tabelas compatível com SQL. Desde o Amazon EMR 6.5.0, você pode usar o Apache Spark 3 em clusters do Amazon EMR com o formato de tabela Iceberg.

A tabela a seguir lista a versão do Iceberg incluída na versão mais recente da série 7.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Iceberg.

Para a versão dos componentes instalados com o Iceberg nesta versão, consulte Versões de componentes da versão 7.9.0.

Informações sobre a versão do Iceberg para o emr-7.9.0

Rótulo de versão do Amazon EMR	Versão do Iceberg	Componentes instalados com o Iceberg
emr-7.9.0	Iceberg 1.7.1-amzn-2	Not available.

A tabela a seguir lista a versão do Iceberg incluída na versão mais recente da série 6.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Iceberg.

Para obter a versão dos componentes instalados com o Iceberg nessa versão, consulte Release 6.15.0 Component Versions.

Informações de versão do Iceberg para o emr-6.15.0

Rótulo de versão do Amazon EMR	Versão do Iceberg	Componentes instalados com o Iceberg
emr-6.15.0	Iceberg 1.4.0-amzn-0	Not available.

Tópicos

- Como o Iceberg funciona
- Usar um cluster com o Iceberg

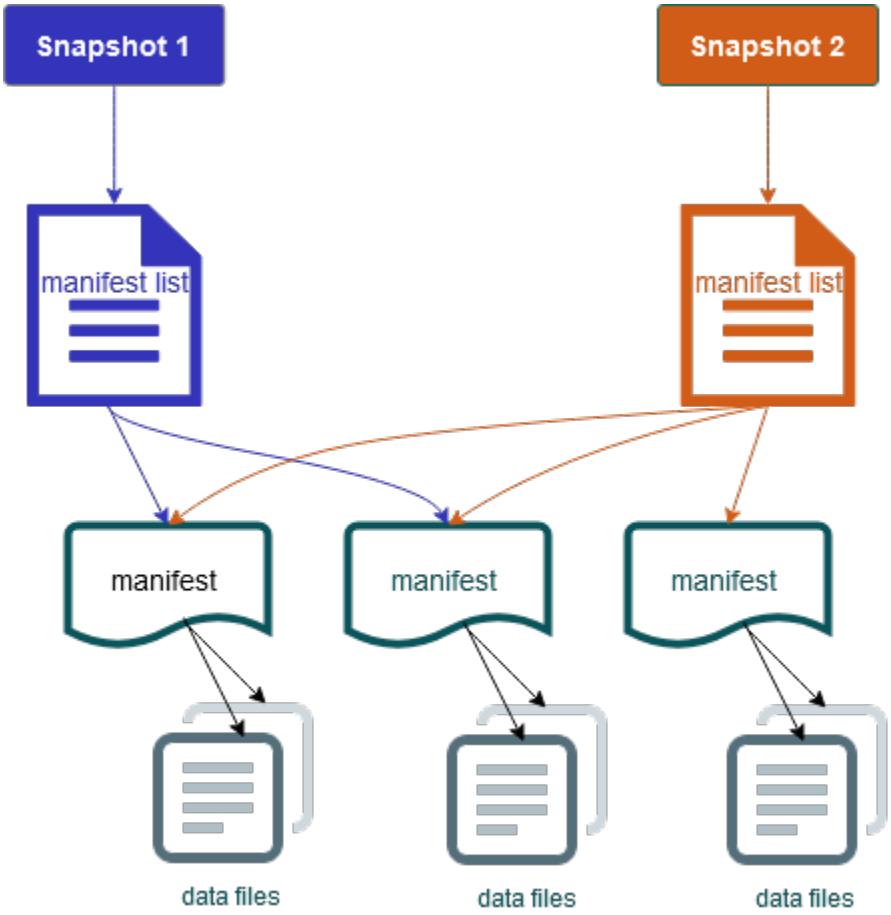
- Histórico de versões do Iceberg

Como o Iceberg funciona

O Iceberg rastreia arquivos de dados individuais em uma tabela em vez de em diretórios. Dessa forma, os gravadores podem criar arquivos de dados no local (os arquivos não são movidos nem alterados). Além disso, os gravadores só podem adicionar arquivos à tabela em uma confirmação explícita. O estado da tabela é mantido em arquivos de metadados. Todas as alterações no estado da tabela criam um novo arquivo de metadados que substitui atomicamente os metadados antigos. O arquivo de metadados da tabela rastreia o esquema da tabela, a configuração do particionamento e outras propriedades.

Ele também inclui snapshots do conteúdo da tabela. Cada snapshot é um conjunto completo de arquivos de dados na tabela em um momento específico. Os snapshots são listados no arquivo de metadados, mas os arquivos de um snapshot são armazenados em arquivos manifesto separados. As transições atômicas de um arquivo de metadados da tabela para o próximo fornecem isolamento de snapshots. Os leitores usam o instantâneo que estava atualizado quando carregaram os metadados da tabela. Os leitores não são afetados pelas alterações até atualizarem e escolherem um novo local de metadados. Arquivos de dados em snapshots são armazenados em um ou mais arquivos manifesto que contêm uma linha para cada arquivo de dados na tabela, seus dados de partição e suas métricas. Um snapshot é a união de todos os arquivos nos seus manifestos. Os arquivos manifesto também podem ser compartilhados entre snapshots para evitar a regravação de metadados que são alterados com pouca frequência.

Diagrama de snapshots do Iceberg



O Iceberg oferece os seguintes atributos:

- Oferece suporte a transações ACID e viagens no tempo no data lake do Amazon S3.
- As novas tentativas de confirmação se beneficiam das vantagens de performance da simultaneidade otimista.
- A resolução de conflitos em nível de arquivo resulta em alta simultaneidade.
- Com estatísticas mínimas e máximas por coluna nos metadados, você pode pular arquivos, o que aumenta a performance de consultas seletivas.
- Você pode organizar tabelas em layouts de partição flexíveis, com a evolução da partição permitindo atualizações nos esquemas de partição. As consultas e os volumes de dados podem, então, mudar sem contar com diretórios físicos.
- Oferece suporte à evolução e à aplicação do esquema.
- As tabelas do Iceberg funcionam como coletores idempotentes e fontes reproduzíveis. Isso permite streaming e suporte em lote com pipelines exatamente uma vez. Os coletores idempotentes rastreiam operações de gravação que foram bem-sucedidas no passado. Portanto, o coletor pode

solicitar dados novamente em caso de falha e descartar dados que tiverem sido enviados várias vezes.

- Visualize o histórico e a linhagem, incluindo evolução da tabela, o histórico de operações e as estatísticas de cada confirmação.
- Migre de um conjunto de dados existente com uma opção de formato de dados (Parquet, ORC, Avro) e de um mecanismo de analytics (Spark, Trino, PrestoDB, Flink, Hive).

Usar um cluster com o Iceberg

Esta seção inclui informações sobre como usar o Iceberg com Spark, Trino, Flink e Hive.

Usar um cluster do Iceberg com o Spark

A partir da versão 6.5.0 do Amazon EMR, é possível usar o Iceberg com o cluster do Spark sem a necessidade de incluir ações de bootstrap. Nas versões 6.4.0 e anteriores do Amazon EMR, é possível usar uma ação de bootstrap para pré-instalar todas as dependências necessárias.

Neste tutorial, você usa o AWS CLI para trabalhar com o Iceberg em um cluster do Amazon EMR Spark. Para usar o console para criar um cluster com o Iceberg instalado, siga as etapas em Criar um data lake no Apache Iceberg usando o Amazon Athena, o Amazon EMR e o AWS Glue.

Criar um cluster do Iceberg

Você pode criar um cluster com o Iceberg instalado usando o AWS Management Console, o AWS CLI ou a API do Amazon EMR. Neste tutorial, você usa o AWS CLI para trabalhar com o Iceberg em um cluster do Amazon EMR. Para usar o console para criar um cluster com o Iceberg instalado, siga as etapas em Criar um data lake no Apache Iceberg usando o Amazon Athena, o Amazon EMR e o AWS Glue.

Para usar o Iceberg no Amazon EMR com AWS CLI o, primeiro crie um cluster com as etapas a seguir. Para obter informações sobre como especificar a classificação do Iceberg usando o AWS CLI, consulte Fornecer uma configuração usando a AWS CLI ao criar um cluster ou. Fornecer uma configuração usando o SDK do Java ao criar um cluster

1. Crie um arquivo `configurations.json`, com o seguinte conteúdo:

```
[{
    "Classification":"iceberg-defaults",
```

```
    "Properties":{"iceberg.enabled":"true"}
}]
```

2. Em seguida, crie um cluster com a configuração a seguir. Substitua o exemplo do caminho do bucket do Amazon S3 e o ID da sub-rede pelos seus.

```
aws emr create-cluster --release-label emr-6.5.0 \
--applications Name=Spark \
--configurations file://configurations.json \
--region us-east-1 \
--name My_Spark_Iceberg_Cluster \
--log-uri s3://amzn-s3-demo-bucket/ \
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole_V2 \
--ec2-attributes
 InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef0
```

Como alternativa, é possível criar um cluster do Amazon EMR que inclua a aplicação do Spark e o arquivo `/usr/share/aws/iceberg/lib/iceberg-spark3-runtime.jar` como uma dependência do JAR em um trabalho do Spark. Para obter mais informações, consulte Submitting Applications.

Para incluir o jar como uma dependência em um trabalho do Spark, adicione a seguinte propriedade de configuração à aplicação do Spark:

```
--conf "spark.jars=/usr/share/aws/iceberg/lib/iceberg-spark3-runtime.jar"
```

Para obter mais informações sobre as dependências de trabalho do Spark, consulte Gerenciamento de dependências no documento do Apache Spark Executar o Spark no Kubernetes.

Inicializar uma sessão do Spark para Iceberg

Os exemplos a seguir demonstram como iniciar o shell interativo do Spark, usar o envio do Spark ou usar os Cadernos do Amazon EMR para trabalhar com o Iceberg no Amazon EMR.

spark-shell

1. Conecte-se ao nó principal usando SSH. Para obter mais informações, consulte Conectar-se ao nó principal usando SSH no Guia de gerenciamento do Amazon EMR.

2. Digite o seguinte comando para iniciar o shell do Spark. Para usar a PySpark concha, `spark-shell` substitua por`pyspark`.

```
spark-shell \
    --conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
    --conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket/prefix/
    --conf spark.sql.catalog.my_catalog.type=glue \
    --conf spark.sql.defaultCatalog=my_catalog \
    --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

spark-submit

1. Conecte-se ao nó principal usando SSH. Para obter mais informações, consulte Conectar-se ao nó principal usando SSH no Guia de gerenciamento do Amazon EMR.
2. Insira o comando a seguir para iniciar a sessão do Spark no Iceberg.

```
spark-submit \
--conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
--conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket1/prefix \
--conf spark.sql.catalog.my_catalog.type=glue \
--conf spark.sql.defaultCatalog=my_catalog \
--conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

EMR Studio notebooks

Para inicializar uma sessão do Spark usando Cadernos do EMR Studio, configure a sessão do Spark usando o comando mágico `%%configure` no Caderno do Amazon EMR, como no exemplo a seguir. Para obter mais informações, consulte Use EMR Notebooks magics no Guia de gerenciamento do Amazon EMR.

```
%%configure -f{
"conf":{
    "spark.sql.catalog.my_catalog":"org.apache.iceberg.spark.SparkCatalog",
    "spark.sql.catalog.my_catalog.type":"glue",
    "spark.sql.catalog.my_catalog.warehouse":"s3://amzn-s3-demo-bucket1/prefix/",
    "spark.sql.defaultCatalog", "my_catalog",
```

```
 "spark.sql.extensions":"org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions"
    }
}
```

CLI

Para inicializar um cluster do Spark usando a CLI e definir todas as configurações padrão da sessão do Spark Iceberg, execute o exemplo a seguir. Para obter mais informações sobre como especificar uma classificação de configuração usando a AWS CLI API do Amazon EMR, consulte Configurar aplicativos.

```
[
  {
    "Classification": "spark-defaults",
    "Properties": {
      "spark.sql.catalog.my_catalog":"org.apache.iceberg.spark.SparkCatalog",
      "spark.sql.catalog.my_catalog.type":"glue",
      "spark.sql.catalog.my_catalog.warehouse":"s3://amzn-s3-demo-bucket1/prefix/",
      "spark.sql.defaultCatalog", "my_catalog",

 "spark.sql.extensions":"org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions"
    }
  }
]
```

Gravar em uma tabela do Iceberg

O exemplo a seguir mostra como criar um DataFrame e gravá-lo como um conjunto de dados do Iceberg. Os exemplos demonstram como trabalhar com conjuntos de dados usando o shell do Spark durante a conexão com o nó principal usando SSH como usuário padrão do hadoop.

> **Note**
>
> Para colar exemplos de código no shell do Spark, digite `:paste` no prompt, cole o exemplo e pressione `CTRL+D`.

PySpark

O Spark inclui um shell baseado em Python, `pyspark`, que você pode usar para gerar protótipos de programas Spark escritos em Python. Invoque `pyspark` no nó principal.

```
## Create a DataFrame.
data = spark.createDataFrame([
 ("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
 ("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
 ("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
 ("103", "2015-01-01", "2015-01-01T13:51:40.519832Z")
],["id", "creation_date", "last_update_time"])

## Write a DataFrame as a Iceberg dataset to the Amazon S3 location.
spark.sql("""CREATE TABLE IF NOT EXISTS dev.db.iceberg_table (id string,
creation_date string,
last_update_time string)
USING iceberg
location 's3://amzn-s3-demo-bucket/example-prefix/db/iceberg_table'""")

data.writeTo("dev.db.iceberg_table").append()
```

Scala

```
import org.apache.spark.sql.SaveMode
import org.apache.spark.sql.functions._

// Create a DataFrame.
val data = Seq(
("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
("103", "2015-01-01", "2015-01-01T13:51:40.519832Z")
).toDF("id", "creation_date", "last_update_time")

// Write a DataFrame as a Iceberg dataset to the Amazon S3 location.
spark.sql("""CREATE TABLE IF NOT EXISTS dev.db.iceberg_table (id string,
creation_date string,
last_update_time string)
USING iceberg
location 's3://amzn-s3-demo-bucket/example-prefix/db/iceberg_table'""")

data.writeTo("dev.db.iceberg_table").append()
```

Ler em uma tabela do Iceberg

PySpark

```
df = spark.read.format("iceberg").load("dev.db.iceberg_table")
df.show()
```

Scala

```
val df = spark.read.format("iceberg").load("dev.db.iceberg_table")
df.show()
```

Spark SQL

```
SELECT * from dev.db.iceberg_table LIMIT 10
```

Usando o AWS Glue Data Catalog com o Spark Iceberg

Você pode se conectar ao AWS Glue Data Catalog a partir do Spark Iceberg. Esta seção mostra diferentes comandos para conexão.

Conecte-se ao catálogo padrão do AWS Glue em sua região padrão

Este exemplo mostra como se conectar usando o tipo de catálogo Glue. Se você não especificar uma ID de catálogo, ela usará o padrão:

```
spark-submit \
    --conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
    --conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket1/prefix \
    --conf spark.sql.catalog.my_catalog.type=glue \
    --conf spark.sql.defaultCatalog=my_catalog \
    --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

Conecte-se a um catálogo do AWS Glue com um ID de catálogo específico

Este exemplo mostra como se conectar usando uma ID de catálogo:

```
spark-submit \
    --conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
    --conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket1/prefix \
```

```
    --conf spark.sql.catalog.my_catalog.type=glue \
    --conf spark.sql.catalog.my_catalog.glue.id=AWS Glue catalog ID \
    --conf spark.sql.defaultCatalog=my_catalog \
    --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

Esse comando pode ser usado para se conectar a um catálogo AWS Glue em uma conta diferente, a um catálogo RMS ou a um catálogo federado.

Usando o Catálogo REST do Iceberg (IRC) com o Spark Iceberg

As seções a seguir detalham como configurar a integração do Iceberg com um catálogo.

Conecte-se ao AWS endpoint IRC do Glue Data Catalog

Veja a seguir um exemplo de `spark-submit` comando para usar o Iceberg REST:

```
spark-submit \
    --conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
    --conf spark.sql.catalog.my_catalog.warehouse=glue catalog ID \
    --conf spark.sql.catalog.my_catalog.type=rest \
    --conf spark.sql.catalog.my_catalog.uri=glue endpoint URI/iceberg \
    --conf spark.sql.catalog.my_catalog.rest.sigv4-enabled=true \
    --conf spark.sql.catalog.my_catalog.rest.signing-name=glue \
    --conf spark.sql.defaultCatalog=my_catalog \
    --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

Para usá-lo em um cluster habilitado para função de tempo de execução, as seguintes configurações adicionais do Spark são necessárias:

```
"spark.hadoop.fs.s3.credentialsResolverClass":
 "software.amazon.glue.GlueTableCredentialsResolver",
"spark.hadoop.catalog-impl": "org.apache.iceberg.aws.glue.GlueCatalog",
"spark.hadoop.glue.id": glue catalog ID
"spark.hadoop.glue.endpoint": "glue endpoint"
```

Para ver a lista de URLs de endpoints do AWS Glue para cada região, consulte Pontos de extremidade e AWS cotas do Glue.

Conecte-se a um endpoint de IRC arbitrário

Veja a seguir um exemplo de `spark-submit` comando para usar um endpoint de IRC:

```
spark-submit \
    --conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
    --conf spark.sql.catalog.my_catalog.warehouse=warehouse name \
    --conf spark.sql.catalog.my_catalog.type=rest \
    --conf spark.sql.catalog.my_catalog.uri=your rest endpoint \
    --conf spark.sql.defaultCatalog=my_catalog \
    --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

Diferenças de configuração quando você usa o Iceberg versus SparkCatalog SparkSessionCatalog

O Iceberg disponibiliza duas maneiras de criar catálogos do Spark Iceberg. Você pode definir a configuração do Spark para uma `SparkCatalog` ou para`SparkSessionCatalog`.

Usando o Iceberg SparkCatalog

A seguir, é mostrado o comando para usar SparkCatalogcomo catálogo do Spark Iceberg:

```
spark-shell \
--conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
--conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket1/prefix \
--conf spark.sql.catalog.my_catalog.type=glue \
--conf spark.sql.defaultCatalog=my_catalog
```

Considerações sobre essa abordagem:

- Você pode acessar as mesas Iceberg, mas nenhuma outra mesa.
- O nome do catálogo não pode ser spark_catalog. Esse é o nome do catálogo inicial no Spark. Ele sempre se conecta a uma metastore do Hive. É o catálogo padrão no Spark, a menos que o usuário o substitua usando. `spark.sql.defaultCatalog`
- Você pode `spark.sql.defaultCatalog` definir o nome do seu catálogo para torná-lo o catálogo padrão.

Usando o Iceberg SparkSessionCatalog

A seguir, é mostrado o comando para usar SparkSessionCatalogcomo catálogo do Spark Iceberg:

```
spark-shell \
```

```
    --conf spark.sql.catalog.spark_catalog=org.apache.iceberg.spark.SparkSessionCatalog
\
    --conf spark.sql.catalog.spark_catalog.warehouse=s3://amzn-s3-demo-bucket1/prefix \
    --conf spark.sql.catalog.spark_catalog.type=glue
```

Considerações sobre essa abordagem:

- Se uma tabela não for encontrada como uma tabela Iceberg, o Spark tentará ver se é uma tabela na metastore do Hive. Consulte Usando o AWS Glue Data Catalog como catálogo do Hive para obter mais informações.
- O nome do catálogo deve ser spark_catalog.

Usando extensões do Iceberg Spark

O Iceberg oferece a extensão Spark `org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions` que os usuários podem configurar por meio da configuração das extensões do Spark. `spark.sql.extensions` As extensões habilitam os principais recursos do Iceberg, como DELETE, UPDATE e MERGE em nível de linha, declarações e procedimentos de linguagem de definição de dados do Spark específicos do Iceberg, como compactação, expiração de instantâneos, ramificação e marcação, etc. Consulte o seguinte para obter mais detalhes:

- Extensões de gravação do Iceberg Spark: Spark Writes
- Extensões DDL do Iceberg Spark: extensões ALTER TABLE SQL
- Extensões do procedimento do Iceberg Spark: Procedimentos Spark

Considerações sobre o uso do Iceberg com o Spark

- Por padrão, o Amazon EMR 6.5.0 não é compatível com a execução do Iceberg no Amazon EMR no EKS. Uma imagem personalizada do Amazon EMR 6.5.0 está disponível para que você possa passar `--jars local:///usr/share/aws/iceberg/lib/iceberg-spark3-runtime.jar` como parâmetro `spark-submit` para criar tabelas do Iceberg no Amazon EMR no EKS. Para obter mais informações, consulte Enviar uma workload do Spark no Amazon EMR usando uma imagem personalizada no Catálogo de desenvolvimento do Amazon EMR no EKS. Você também pode entrar em contato com Suporte para obter assistência. Desde o Amazon EMR 6.6.0, o Iceberg é compatível com o Amazon EMR no EKS.

- Ao usar o AWS Glue como um catálogo para o Iceberg, certifique-se de que o banco de dados no qual você está criando uma tabela exista no AWS Glue. Se você estiver usando serviços como AWS Lake Formation e não conseguir carregar o catálogo, verifique se você tem acesso adequado ao serviço para executar o comando.
- Se você usa o Iceberg SparkSessionCatalog, conforme descrito emDiferenças de configuração quando você usa o Iceberg versus SparkCatalog SparkSessionCatalog, você deve seguir as etapas de configuração descritas em Configurar o catálogo de dados do AWS Glue como metastore do Apache Hive, além de definir as configurações do catálogo de dados do Spark Iceberg Glue. AWS

Usar um cluster do Iceberg com o Trino

Desde a versão 6.6.0 do Amazon EMR, você pode usar o Iceberg com o cluster do Trino.

Neste tutorial, você usa o AWS CLI para trabalhar com o Iceberg em um cluster Trino do Amazon EMR. Para usar o console para criar um cluster com o Iceberg instalado, siga as etapas em Criar um data lake no Apache Iceberg usando o Amazon Athena, o Amazon EMR e o AWS Glue.

Criar um cluster do Iceberg

Para usar o Iceberg no Amazon EMR com AWS CLI o, primeiro crie um cluster com as etapas a seguir. Para obter informações sobre como especificar a classificação do Iceberg usando o AWS CLI, consulte Fornecer uma configuração usando a AWS CLI ao criar um cluster ou. Fornecer uma configuração usando o SDK do Java ao criar um cluster

1. Crie um arquivo `iceberg.properties` e defina um valor para o catálogo escolhido. Por exemplo, se você quiser usar a metastore do Hive como catálogo, o arquivo deverá ter o conteúdo a seguir.

   ```
   connector.name=iceberg
   hive.metastore.uri=thrift://localhost:9083
   ```

   Se você quiser usar o AWS Glue Data Catalog como sua loja, seu arquivo deve ter o seguinte conteúdo.

   ```
   connector.name=iceberg
   iceberg.catalog.type=glue
   ```

2. Crie uma ação de bootstrap que copie `iceberg.properties` do Amazon S3 para `/etc/trino/conf/catalog/iceberg.properties`, como no exemplo a seguir. Para obter mais informações sobre ações de bootstrap, consulte Criar ações de bootstrap para instalar softwares adicionais.

```
set -ex
sudo aws s3 cp s3://amzn-s3-demo-bucket/iceberg.properties /etc/trino/conf/catalog/
iceberg.properties
```

3. Crie um cluster com a configuração a seguir, substituindo o exemplo de caminho de script de ações de bootstrap e o nome da chave pelos seus.

```
aws emr create-cluster --release-label emr-6.7.0 \
--applications Name=Trino \
--region us-east-1 \
--name My_Trino_Iceberg_Cluster \
--bootstrap-actions '[{"Path":"s3://amzn-s3-demo-bucket","Name":"Add
 iceberg.properties"}]' \
--instance-groups InstanceGroupType=MASTER,InstanceCount=1,InstanceType=c3.4xlarge
 InstanceGroupType=CORE,InstanceCount=3,InstanceType=c3.4xlarge \
--use-default-roles \
--ec2-attributes KeyName=<key-name>
```

Inicializar uma sessão do Trino para Iceberg

Para inicializar uma sessão no Trino, execute o comando a seguir.

```
trino-cli --catalog iceberg
```

Gravar em uma tabela do Iceberg

Crie e grave na tabela com os comandos SQL a seguir.

```
trino> SHOW SCHEMAS;
trino> CREATE TABLE default.iceberg_table (
            id int,
            data varchar,
            category varchar)
       WITH (
            format = 'PARQUET',
```

```
            partitioning = ARRAY['category', 'bucket(id, 16)'],
            location = 's3://amzn-s3-demo-bucket/<prefix>')

trino> INSERT INTO default.iceberg_table VALUES (1,'a','c1'), (2,'b','c2'),
 (3,'c','c3');
```

Ler em uma tabela do Iceberg

Para ler a tabela do Iceberg, execute o comando a seguir.

```
trino> SELECT * from default.iceberg_table;
```

Considerações sobre o uso do Iceberg com o Trino

- O Amazon EMR 6.5 não é nativamente compatível com o catálogo do Iceberg com o Trino. O Trino precisa do Iceberg v0.11, por isso recomendamos iniciar um cluster do Amazon EMR para Trino separado do cluster do Spark e incluir o Iceberg v0.11 neste cluster.
- Ao usar o AWS Glue como um catálogo para o Iceberg, certifique-se de que o banco de dados no qual você está criando uma tabela exista no AWS Glue. Se você estiver usando serviços como AWS Lake Formation e não conseguir carregar o catálogo, verifique se você tem acesso adequado ao serviço para executar o comando.
- A integração com o Iceberg Glue não funciona com o catálogo de armazenamento gerenciado do Redshift.

Usar um cluster do Iceberg com o Flink

Desde a versão 6.9.0 do Amazon EMR, você pode usar o Iceberg com um cluster do Flink sem as etapas de configuração necessárias ao usar a integração Iceberg-Flink de código aberto.

Criar um cluster no Iceberg

É possível criar um cluster com o Iceberg instalado usando o AWS Management Console, a AWS CLI ou a API do Amazon EMR. Neste tutorial, você usa o AWS CLI para trabalhar com o Iceberg em um cluster do Amazon EMR. Para usar o console para criar um cluster com o Iceberg instalado, siga as etapas em Criar um data lake no Apache Iceberg usando o Amazon Athena, o Amazon EMR e o AWS Glue.

Para usar o Iceberg no Amazon EMR com AWS CLI o, primeiro crie um cluster com as etapas a seguir. Para obter informações sobre como especificar a classificação do Iceberg usando o AWS

CLI, consulte Fornecer uma configuração usando a AWS CLI ao criar um cluster ou. Fornecer uma configuração usando o SDK do Java ao criar um cluster Crie um arquivo denominado `configurations.json` com o seguinte conteúdo:

```
[{
"Classification":"iceberg-defaults",
    "Properties":{"iceberg.enabled":"true"}
}]
```

Em seguida, crie um cluster com a configuração a seguir, substituindo o exemplo de caminho do bucket do Amazon S3 e o ID da sub-rede pelos seus próprios valores:

```
aws emr create-cluster --release-label emr-6.9.0 \
--applications Name=Flink \
--configurations file://iceberg_configurations.json \
--region us-east-1 \
--name My_flink_Iceberg_Cluster \
--log-uri s3://amzn-s3-demo-bucket/ \
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole \
--ec2-attributes InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef
```

Como alternativa, é possível criar um cluster do Amazon EMR 6.9.0 contendo uma aplicação do Flink e usar o arquivo `/usr/share/aws/iceberg/lib/iceberg-flink-runtime.jar` como uma dependência do JAR em um trabalho do Flink.

Usar o clinte SQL no Flink

O script do cliente SQL está localizado em `/usr/lib/flink/bin`. Você pode executar o script com o seguinte comando:

```
flink-yarn-session -d # starting the Flink YARN Session in detached mode
./sql-client.sh
```

Isso inicia um shell SQL no Flink.

Exemplos do Flink

Criar uma tabela no Iceberg

SQL no Flink

```
CREATE CATALOG glue_catalog WITH (
   'type'='iceberg',
   'warehouse'='<WAREHOUSE>',
   'catalog-type'='glue'
 );

USE CATALOG  glue_catalog;

CREATE DATABASE IF NOT EXISTS <DB>;

USE <DB>;

CREATE TABLE IF NOT EXISTS `glue_catalog`.`<DB>`.`sample` (id int, data string);
```

API de tabela

```
EnvironmentSettings settings =
                EnvironmentSettings.newInstance().inBatchMode().build();

TableEnvironment tEnv = TableEnvironment.create(settings);

String warehouse = "<WAREHOUSE>";
String db = "<DB>";

tEnv.executeSql(
                "CREATE CATALOG glue_catalog WITH (\n"
                        + "   'type'='iceberg',\n"
                        + "   'warehouse'='"
                        + warehouse
                        + "',\n"
                        + "   'catalog-impl'='org.apache.iceberg.aws.glue.GlueCatalog',
\n"
                        + "   'io-impl'='org.apache.iceberg.aws.s3.S3FileIO'\n"
                        + " );");

tEnv.executeSql("USE CATALOG  glue_catalog;");
tEnv.executeSql("CREATE DATABASE IF NOT EXISTS " + db + ";");
```

```
tEnv.executeSql("USE " + db + ";");
tEnv.executeSql(
        "CREATE TABLE `glue_catalog`.`" + db + "`.`sample` (id bigint, data string);");
```

Gravar em uma tabela do Iceberg

SQL no Flink

```
INSERT INTO `glue_catalog`.`<DB>`.`sample` values (1, 'a'),(2,'b'),(3,'c');
```

API de tabela

```
tEnv.executeSql(
        "INSERT INTO `glue_catalog`.`"
                + db
                + "`.`sample` values (1, 'a'),(2,'b'),(3,'c');");
```

API de fluxo de dados

```
final StreamExecutionEnvironment env =
 StreamExecutionEnvironment.getExecutionEnvironment();

final StreamTableEnvironment tableEnv = StreamTableEnvironment.create(env);

String db = "<DB Name>";

String warehouse = "<Warehouse Path>";

GenericRowData rowData1 = new GenericRowData(2);
rowData1.setField(0, 1L);
rowData1.setField(1, StringData.fromString("a"));

DataStream<RowData> input = env.fromElements(rowData1);

Map<String, String> props = new HashMap<();
props.put("type", "iceberg");
props.put("warehouse", warehouse);
props.put("io-impl", "org.apache.iceberg.aws.s3.S3FileIO");

CatalogLoader glueCatlogLoader =
        CatalogLoader.custom(
                "glue",
```

```
                props,
                new Configuration(),
                "org.apache.iceberg.aws.glue.GlueCatalog");

TableLoader tableLoader =
        TableLoader.fromCatalog(glueCatlogLoader, TableIdentifier.of(db, "sample"));

DataStreamSink<Void> dataStreamSink =
        FlinkSink.forRowData(input).tableLoader(tableLoader).append();

env.execute("Datastream Write");
```

Ler em uma tabela do Iceberg

SQL no Flink

```
SELECT * FROM `glue_catalog`.`<DB>`.`sample`;
```

API de tabela

```
Table result = tEnv.sqlQuery("select * from `glue_catalog`.`" + db + "`.`sample`;");
```

API de fluxo de dados

```
final StreamExecutionEnvironment env =
 StreamExecutionEnvironment.getExecutionEnvironment();

final StreamTableEnvironment tableEnv = StreamTableEnvironment.create(env);

String db = "<DB Name>";

String warehouse = "<Warehouse Path>";

Map<String, String> props = new HashMap<>();
props.put("type", "iceberg");
props.put("warehouse", warehouse);
props.put("io-impl", "org.apache.iceberg.aws.s3.S3FileIO");

CatalogLoader glueCatlogLoader =
        CatalogLoader.custom(
                "glue",
                props,
```

```
                new Configuration(),
                "org.apache.iceberg.aws.glue.GlueCatalog");

TableLoader tableLoader =
        TableLoader.fromCatalog(glueCatlogLoader, TableIdentifier.of(db, "sample"));

DataStream<RowData> batch =

 FlinkSource.forRowData().env(env).tableLoader(tableLoader).streaming(false).build();

batch.print().name("print-sink");
```

Usar o catálogo do Hive

Certifique-se de que as dependências do Flink e do Hive sejam resolvidas conforme descrito em Configurar o Flink com o Hive Metastore e o Catálogo do Glue.

Executar um trabalho do Flink

Uma forma de enviar um trabalho ao Flink é usar uma sessão do YARN do Flink por trabalho. Isso pode ser iniciado com o seguinte comando:

```
sudo flink run -m yarn-cluster -p 4 -yjm 1024m -ytm 4096m $JAR_FILE_NAME
```

Considerações sobre o uso do Iceberg com o Flink

- Ao usar o AWS Glue como um catálogo para o Iceberg, certifique-se de que o banco de dados no qual você está criando uma tabela exista no AWS Glue. Se você estiver usando serviços como AWS Lake Formation e não conseguir carregar o catálogo, verifique se você tem acesso adequado ao serviço para executar o comando.
- A integração com o Iceberg Glue não funciona com o catálogo de armazenamento gerenciado do Redshift.

Usar um cluster do Iceberg com o Hive

Com as versões 6.9.0 e posteriores do Amazon EMR, você pode usar o Iceberg com um cluster do Hive sem precisar executar as etapas de configuração necessárias para a integração do IcebergHive de código aberto. Para as versões 6.8.0 e anteriores do Amazon EMR, você pode usar uma ação de bootstrap para instalar o jar `iceberg-hive-runtime` e configurar o suporte do Hive para Iceberg.

O Amazon EMR 6.9.0 inclui todos os atributos para a integração do Hive 3.1.3 com o Iceberg 0.14.1 e também inclui atributos adicionais do Amazon EMR, como seleção automática de mecanismos de execução compatíveis em runtime (Amazon EMR no EKS 6.9.0).

Criar um cluster do Iceberg

Você pode criar um cluster com o Iceberg instalado usando o AWS Management Console, o AWS CLI ou a API do Amazon EMR. Neste tutorial, você usa o AWS CLI para trabalhar com o Iceberg em um cluster do Amazon EMR. Para usar o console para criar um cluster com o Iceberg instalado, siga as etapas em Criar um data lake no Iceberg usando o Amazon Athena, o Amazon EMR e o AWS Glue.

Para usar o Iceberg no Amazon EMR com AWS CLI o, primeiro crie um cluster usando as etapas abaixo. Para obter informações sobre como especificar a classificação Iceberg usando o AWS CLI ou o Java SDK, consulte ou. Fornecer uma configuração usando a AWS CLI ao criar um cluster Fornecer uma configuração usando o SDK do Java ao criar um cluster Crie um arquivo chamado `configurations.json` com o seguinte conteúdo:

```
[{
    "Classification":"iceberg-defaults",
    "Properties":{"iceberg.enabled":"true"}
}]
```

Em seguida, crie um cluster com a configuração a seguir, substituindo o exemplo de caminho do bucket do Amazon S3 e o ID da sub-rede pelos seus próprios:

```
aws emr create-cluster --release-label emr-6.9.0 \
--applications Name=Hive \
--configurations file://iceberg_configurations.json \
--region us-east-1 \
--name My_hive_Iceberg_Cluster \
--log-uri s3://amzn-s3-demo-bucket/ \
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole \
--ec2-attributes InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef
```

Um cluster do Iceberg no Hive faz o seguinte:

- Carrega o jar de runtime do Hive no Iceberg no Hive e habilita a configuração relacionada ao Iceberg para o mecanismo do Hive.

- Permite a seleção dinâmica do mecanismo de execução do Hive no Amazon EMR para impedir que os usuários configurem um mecanismo de execução compatível com o Iceberg.

> **Note**
>
> Atualmente, os clusters Hive Iceberg não são compatíveis com o AWS Glue Data Catalog. O catálogo padrão do Iceberg é `HiveCatalog`, que corresponde ao metastore configurado para o ambiente Hive. Para obter mais informações sobre gerenciamento de catálogos, consulte Usando HCatalog na documentação do Apache Hive.

Suporte a recursos

O Amazon EMR versão 6.9.0 é compatível com o Hive 3.1.3 e o Iceberg 0.14.1. O suporte a atributos é limitado aos atributos compatíveis com o Iceberg para o Hive 3.1.2 e 3.1.3. Os seguintes comandos são compatíveis:

- Com as versões 6.9.0 a 6.12.x do Amazon EMR, você deve incluir o jar `libfb303` no diretório `auxlib` do Hive. Use o seguinte comando para instalar:

  ```
  sudo /usr/bin/ln -sf /usr/lib/hive/lib/libfb303-*.jar /usr/lib/hive/auxlib/
  libfb303.jar
  ```

  Com as versões 6.13 e posteriores do Amazon EMR, o jar `libfb303` é automaticamente vinculado ao diretório `auxlib` do Hive.

- Criar uma tabela
  - Tabela não particionada: as tabelas externas no Hive podem ser criadas ao ser fornecido o manipulador de armazenamento da seguinte forma:

    ```
    CREATE EXTERNAL TABLE x (i int) STORED BY
     'org.apache.iceberg.mr.hive.HiveIcebergStorageHandler'
    ```

  - Tabela particionada: as tabelas externas particionadas no Hive podem ser criadas da seguinte forma:

    ```
    CREATE EXTERNAL TABLE x (i int) PARTITIONED BY (j int) STORED BY
     'org.apache.iceberg.mr.hive.HiveIcebergStorageHandler'
    ```

> **Note**
>
> O formato de `STORED AS` arquivo do não ORC/AVRO/PARQUET é suportado no Hive 3. A opção padrão e única é Parquet.

- Descartar uma tabela: o comando `DROP TABLE` é usado para descartar tabelas, como no exemplo a seguir:

```
DROP TABLE [IF EXISTS] table_name [PURGE];
```

- Ler uma tabela: as instruções `SELECT` podem ser usadas para ler tabelas do Iceberg no Hive, como no exemplo a seguir. Os mecanismos de execução compatíveis são MR e Tez.

```
SELECT * FROM table_name
```

  Para obter informações sobre a sintaxe de seleção do Hive, consulte LanguageManual Selecionar. Para obter informações sobre instruções selecionadas com tabelas do Iceberg no Hive, consulte Selecionar Apache Iceberg.

- Inserir em uma tabela: a instrução `INSERT INTO` do HiveQL funciona em tabelas do Iceberg compatíveis somente com o mecanismo de execução do Map Reduce. Os usuários do Amazon EMR não precisam definir explicitamente o mecanismo de execução porque o Hive no Amazon EMR seleciona o mecanismo para tabelas do Iceberg no runtime.
  - Inserção de tabela única: exemplo:

```
INSERT INTO table_name VALUES ('a', 1);
INSERT INTO table_name SELECT...;
```

  - Inserção de várias tabelas: a inserção de várias tabelas não atômicas em instruções é compatível. Exemplo:

```
FROM source
 INSERT INTO table_1 SELECT a, b
 INSERT INTO table_2 SELECT c,d;
```

A partir do Amazon EMR 7.3.0, o Hive with Iceberg oferece suporte ao AWS Glue Data Catalog como metastore. Para usar o AWS Glue Data Catalog como metastore, defina a seguinte propriedade.

```
SET iceberg.catalog.<catalog_name>.catalog-
impl=org.apache.iceberg.aws.glue.GlueCatalog;
```

Opcionalmente, você também pode definir a propriedade a seguir.

```
SET iceberg.catalog.<catalog_name>.type=glue;
```

Você pode então criar uma tabela usando o exemplo a seguir.

```
CREATE EXTERNAL TABLE table_name (col1 type1, col2 type2,..)
ROW FORMAT SERDE 'org.apache.iceberg.mr.hive.HiveIcebergSerDe'
STORED BY 'org.apache.iceberg.mr.hive.HiveIcebergStorageHandler'
location '<location>'
TBLPROPERTIES ('table_type'='iceberg', 'iceberg.catalog'='<catalog_name>');
```

Considerações sobre o uso do Iceberg com o Hive

- O Iceberg é compatível com os seguintes tipos de consulta:
  - Create table
  - Drop table
  - Inserir na tabela
  - Ler a tabela

- Somente o mecanismo de execução MR (MapReduce) é suportado para operações de DML (linguagem de manipulação de dados), e o MR está obsoleto no Hive 3.1.3.

- AWS Atualmente, o Glue Data Catalog não é compatível com o Iceberg with Hive.

- O tratamento de erros não é suficientemente robusto. Em casos de configuração incorreta, as inserções nas consultas podem ser concluídas com êxito. No entanto, a falha na atualização dos metadados pode resultar em perda de dados.
- A integração com o Iceberg Glue não funciona com o catálogo de armazenamento gerenciado do Redshift.

Histórico de versões do Iceberg

A tabela a seguir lista a versão do Iceberg incluída em cada versão do Amazon EMR, além dos componentes instalados com a aplicação. Para obter as versões dos componentes em cada versão, consulte a seção Versão de componentes da sua versão em Versões 7.x do Amazon EMR, Versões de lançamento 6.x do Amazon EMR ou Versões de lançamento 5.x do Amazon EMR.

Informações sobre a versão do Iceberg

Rótulo de versão do Amazon EMR	Versão do Iceberg	Componentes instalados com o Iceberg
emr-7.9.0	1.7.1-amzn-2	Not available.
emr-7.8.0	1.7.1-amzn-1	Not available.
emr-7.7.0	1.7.1-amzn-0	Not available.
emr-7.6.0	1.6.1-amzn-2	Not available.
emr-7.5.0	1.6.1-amzn-1	Not available.
emr-7.4.0	1.6.1-amzn-0	Not available.
emr-7.3.0	1.5.2-amzn-0	Not available.
emr-7.2.0	1.5.0-amzn-0	Not available.
emr-7.1.0	1.4.3-amzn-0	Not available.
emr-7.0.0	q1.4.2-amzn-0	Not available.
emr-6.15.0	1.4.0-amzn-0	Not available.
emr-6.14.0	1.3.1-amzn-0	Not available.
emr-6.13.0	1.3.0-amzn-1	Not available.
emr-6.12.0	1.3.0-amzn-0	Not available.
emr-6.11.1	1.2.0-amzn-0	Not available.

Rótulo de versão do Amazon EMR	Versão do Iceberg	Componentes instalados com o Iceberg
emr-6.11.0	1.2.0-amzn-0	Not available.
emr-6.10.1	1.1.0-amzn-0	Not available.
emr-6.10.0	1.1.0-amzn-0	Not available.
emr-6.9.1	0.14.1-amzn-0	Not available.
emr-6.9.0	0.14.1-amzn-0	Not available.
emr-6.8.1	0.14.0-amzn-0	Not available.
emr-6.8.0	0.14.0-amzn-0	Not available.
emr-6.7.0	0.13.1-amzn-0	Not available.
emr-6.6.0	0.13.1	Not available.
emr-6.5.0	0.12.0	Not available.

Notas da versão do Iceberg por versão

- Amazon EMR 6.9.0 - Notas da versão do Iceberg

Amazon EMR 6.9.0 - Notas da versão do Iceberg

Amazon EMR 6.9.0 -Alterações no Iceberg

Tipo	Descrição
Recurso	Integração do Flink no Amazon EMR com o Iceberg.
Recurso	Integração do Hive no Amazon EMR com o Iceberg.

Tipo	Descrição
Recurso	Support para armazenar arquivos de metadados do Iceberg no Amazon for Lustre FSx para melhorar o tempo de planejamento da consulta.
Backport	PR 5050: Flink 1.15: suporte para opções de gravação nos comentários SQL de inserção em linha.
Backport	PR 5282 AWS: Corrija falhas de repetição de PUT abrindo novos fluxos de arquivos de dados.
Backport	PR 5318: Flink 1.15: Preenche a lacuna entre FlinkSource e IcebergSource (FLIP-27) e adicionou uma configuração opcional para usar a fonte FLIP-27 no Flink SQL.
Backport	PR 5344: Flink 1.14: Preenche a lacuna entre FlinkSource e IcebergSource (FLIP-27) e adicionou uma configuração opcional para usar a fonte FLIP-27 no Flink SQL.
Backport	PR 5393: Flink 1.14, 1.15: Evite converter métricas Iceberg MetricContext em Flink no leitor de origem FLIP-27.
Backport	PR 5401: Flink 1.14, 1.15: IcebergSo urceReader Grupo perdido no PR #5393 para métricas do leitor de origem do FLIP-27.
Backport	PR 5679: Spark 3.2, 3.3: Corrige a propagação de nulidade para o nó. MergeRows
Backport	PR 5860: Spark 3.3: Correção QueryFailure ao executar RewriteManifestProcedure em tabelas particionadas por data.

Tipo	Descrição
Backport	PR 5880: Spark 3.3: Corrige a nulidade nas projeções. merge-on-read
Backport	PR 5917: Spark 3.2: Corrige a nulidade nas projeções. merge-on-read

Caderno Jupyter no Amazon EMR

O Caderno Jupyter é uma aplicação Web de código aberto que você pode usar para criar e compartilhar documentos que contêm código ativo, equações, visualizações e texto narrativo. O Amazon EMR oferece três opções para trabalhar com cadernos Jupyter:

Tópicos

- EMR Studio
- Caderno do Amazon EMR baseado no Caderno Jupyter
- JupyterHub

EMR Studio

O Amazon EMR Studio é um ambiente de desenvolvimento integrado (IDE) baseado na Web para cadernos Jupyter totalmente gerenciados que são executados em clusters do Amazon EMR. Você pode configurar um EMR Studio para que sua equipe desenvolva, visualize e depure aplicativos escritos em R, Python, Scala e. PySpark

Recomendamos usar o EMR Studio ao usar notebooks Jupyter no Amazon EMR. Para obter mais informações, consulte EMR Studio no Guia de gerenciamento do Amazon EMR.

Caderno do Amazon EMR baseado no Caderno Jupyter

Os Cadernos do EMR constituem um ambiente do Caderno Jupyter incorporado ao console do Amazon EMR que permite criar rapidamente cadernos Jupyter, anexá-los a clusters do Spark e, em seguida, abrir o editor do Caderno Jupyter no console para executar consultas e códigos remotamente. Um caderno do EMR é salvo no Amazon S3 independentemente dos clusters para armazenamento durável, acesso rápido e flexibilidade. Você pode ter vários cadernos abertos, conectar vários cadernos a um único cluster e reutilizar um caderno em clusters diferentes.

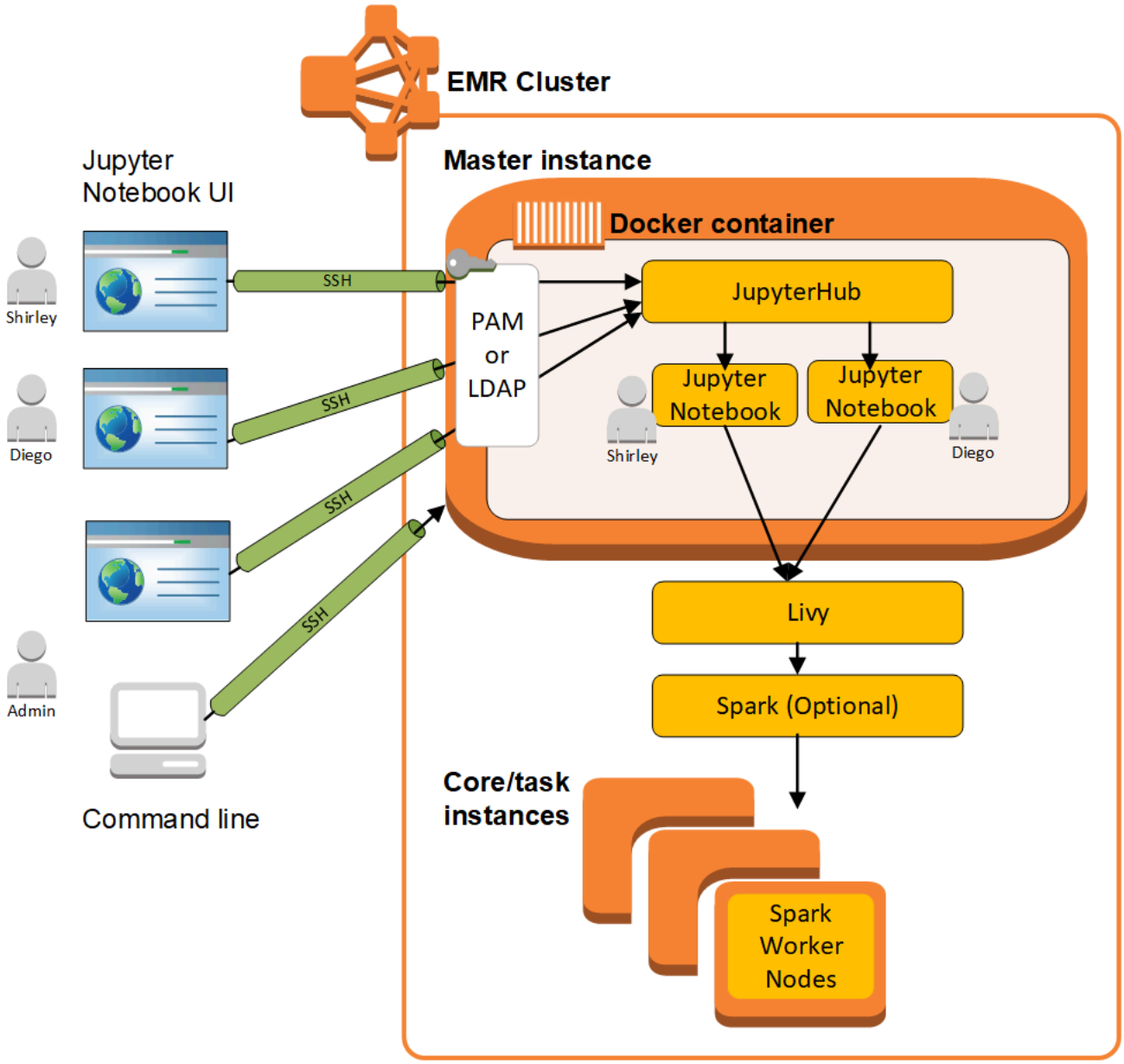
Para obter mais informações, consulte Cadernos do EMR no Guia de gerenciamento do Amazon EMR.

JupyterHub

O Caderno Jupyter é uma aplicação Web de código aberto que você pode usar para criar e compartilhar documentos que contêm código ativo, equações, visualizações e texto narrativo. JupyterHubpermite hospedar várias instâncias de um servidor de caderno Jupyter de usuário único. Quando você cria um cluster com o JupyterHub, o Amazon EMR cria um contêiner do Docker no nó principal do cluster. JupyterHub, todos os componentes necessários para o Jupyter e o Sparkmagic funcionam dentro do contêiner.

O Sparkmagic é uma biblioteca de kernels que permite que os cadernos Jupyter interajam com o Apache Spark em execução no Amazon EMR por meio de Apache Livy, que é um servidor REST para o Spark. O Spark e o Apache Livy são instalados automaticamente quando você cria um cluster com o JupyterHub. O kernel padrão do Python 3 para Jupyter está disponível junto com os kernels PySpark 3,, PySpark, Spark e Spark que estão disponíveis com o Spark. Você pode usar esses kernels para executar o código do Spark ad-hoc e consultas SQL interativas usando o Python e o Scala. Você pode instalar kernels adicionais dentro do contêiner Docker manualmente. Para obter mais informações, consulte Instalar Kernels e bibliotecas adicionais.

O diagrama a seguir descreve os componentes do JupyterHub Amazon EMR com os métodos de autenticação correspondentes para usuários de cadernos e o administrador. Para obter mais informações, consulte Adicionar usuários e administradores do Caderno Jupyter.



A tabela a seguir lista a versão do JupyterHub Amazon EMR 7.x, além dos componentes com os quais o Amazon EMR instala. JupyterHub

Para obter a versão dos componentes instalados com JupyterHub essa versão, consulte Release 7.9.0 9.0

JupyterHub informações de versão do emr-7.9.0

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-7.9.0	JupyterHub 1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

A tabela a seguir lista a versão do JupyterHub incluído na versão mais recente da série 6.x do Amazon EMR, além dos componentes com os quais o Amazon EMR instala. JupyterHub

Para obter a versão dos componentes instalados com JupyterHub essa versão, consulte Release 6.15.0 15.0 15.0

JupyterHub informações de versão do emr-6.15.0

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-6.15.0	JupyterHub 1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
		, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

A tabela a seguir lista a versão JupyterHub incluída na versão mais recente da série 5.x do Amazon EMR, além dos componentes com os quais o Amazon EMR instala. JupyterHub

Para obter a versão dos componentes instalados com JupyterHub essa versão, consulte Release 5.36.2

JupyterHub informações de versão do emr-5.36.2

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-5.36.2	JupyterHub 1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

O kernel do Python 3 incluído no JupyterHub Amazon EMR é 3.6.4.

As bibliotecas instaladas dentro do `jupyterhub` contêiner podem variar entre as versões do Amazon EMR e as versões da EC2 AMI.

Para listar bibliotecas instaladas usando o **`conda`**

- Execute o seguinte comando na linha de comandos do nó principal:

```
sudo docker exec jupyterhub bash -c "conda list"
```

Para listar bibliotecas instaladas usando o **`pip`**

- Execute o seguinte comando na linha de comandos do nó principal:

```
sudo docker exec jupyterhub bash -c "pip freeze"
```

Tópicos

- Criar um cluster com JupyterHub
- Considerações ao usar JupyterHub no Amazon EMR
- Configurando JupyterHub
- Configurar a persistência de cadernos no Amazon S3
- Conectar-se ao nó principal e aos servidores de cadernos
- JupyterHub configuração e administração
- Adicionar usuários e administradores do Caderno Jupyter
- Instalar Kernels e bibliotecas adicionais
- JupyterHub histórico de versões

Criar um cluster com JupyterHub

É possível criar um cluster do Amazon EMR JupyterHub usando o AWS Management Console AWS Command Line Interface, o ou a API do Amazon EMR. Certifique-se de que o cluster não seja criado com a opção para encerrar automaticamente após concluir etapas (opção `--auto-terminate` na AWS CLI). Além disso, certifique-se de que os administradores de notebook e os usuários possam acessar o par de chaves que você usa ao criar o cluster. Para obter mais informações, consulte Usar um par de chaves para credenciais SSH no Guia de gerenciamento do Amazon EMR.

Criar um cluster JupyterHub usando o console

Use o procedimento a seguir para criar um cluster JupyterHub instalado usando Opções avançadas no console do Amazon EMR.

Criar um cluster do Amazon JupyterHub EMR com o

1. Navegue até o novo console do Amazon EMR e selecione Alternar para o console antigo na navegação lateral. Para obter mais informações sobre o que esperar ao alternar para o console antigo, consulte Usar o console antigo.
2. Escolha Create cluster (Criar cluster), Go to advanced options (Ir para opções avançadas).
3. Em Software Configuration (Configuração de software):
   - Em Versão, selecione emr-5.36.2 e escolha. JupyterHub
   - Se você usa o Spark, para usar o Catálogo de Dados do AWS Glue como o metastore para Spark SQL, selecione Usar para metadados da tabela do Spark. Para obter mais informações, consulte Use o catálogo do AWS Glue Data Catalog com o Spark no Amazon EMR.
   - Para Edit software settings (Editar configurações de software), escolha Enter configuration (Inserir configuração) e especifique valores ou escolha Load JSON from S3 (Carregar JSON de S3) e especifique um arquivo de configuração JSON. Para obter mais informações, consulte Configurando JupyterHub.
4. Em (Add steps (optional)) Adicionar etapas (opcional) configure as etapas para serem executadas quando o cluster for criado, certifique-se de que Auto-terminate cluster after the last step is completed (Encerrar o cluster automaticamente após a última etapa for concluída) não esteja selecionada e escolha Next (Próximo).
5. Escolha a opção Hardware Configuration (Configuração de hardware), Next (Próximo). Para obter mais informações, consulte Configurar o hardware e a rede do cluster no Guia de gerenciamento do Amazon EMR.
6. Escolha opções para General Cluster Settings (Configurações gerais do cluster), Next (Próximo).
7. Escolha Security Options (Opões de segurança), especificando um par de chaves e escolha Create Cluster (Criar cluster).

Crie um cluster JupyterHub usando o AWS CLI

Para iniciar um cluster com JupyterHub, use o `aws emr create-cluster` comando e, para a `--applications` opção, especifique`Name=JupyterHub`. O exemplo a seguir inicia um JupyterHub

cluster no Amazon EMR com duas EC2 instâncias (uma instância principal e uma instância core). Além disso, a depuração é habilitada, com os logs armazenados no local do Amazon S3 conforme especificado pelo `--log-uri`. O par de chaves especificado fornece acesso a EC2 instâncias da Amazon no cluster.

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr create-cluster --name="MyJupyterHubCluster" --release-label emr-5.36.2 \
--applications Name=JupyterHub --log-uri s3://amzn-s3-demo-bucket/MyJupyterClusterLogs
 \
--use-default-roles --instance-type m5.xlarge --instance-count 2 --ec2-attributes
 KeyName=MyKeyPair
```

Considerações ao usar JupyterHub no Amazon EMR

Considere o seguinte ao usar JupyterHub no Amazon EMR.

- > **Warning**
  >
  > Os notebooks e arquivos do usuário são salvos no sistema de arquivos no nó principal. Este é um armazenamento temporário que não persiste por meio da finalização do cluster. Quando um cluster é finalizado, esses dados são perdidos se não forem feitos backup. Recomendamos que você programe backups regulares usando trabalhos `cron` ou outro meio adequado para seu aplicativo.
  > Além disso, as alterações de configuração feitas no contêiner podem não persistir se o contêiner for reiniciado. Recomendamos que você faça script ou automatize a configuração de contêiner para que possa reproduzir personalizações com mais facilidade.
- A autenticação Kerberos que foi configurada usando uma configuração de segurança do Amazon EMR não é compatível.
- Não há suporte ao OAuthenticator.

Configurando JupyterHub

Você pode personalizar a configuração do JupyterHub Amazon EMR e cadernos de usuários individuais conectando-se ao nó principal do cluster e editando os arquivos de configuração. Depois de alterar os valores, reinicie o contêiner `jupyterhub`.

Modifique as propriedades dos seguintes arquivos para configurar JupyterHub cadernos Jupyter individuais:

- `jupyterhub_config.py`: por padrão, esse arquivo é salvo no diretório `/etc/jupyter/conf/` no nó principal. Para obter mais informações, consulte Noções básicas de configuração na JupyterHub documentação.
- `jupyter_notebook_config.py`: esse arquivo é salvo no diretório `/etc/jupyter/` por padrão e copiado para o contêiner `jupyterhub` como padrão. Para obter mais informações, consulte Arquivo de configuração e opções de linha de comando na documentação do Notebook Jupyter.

Você também pode usar a classificação de configuração `jupyter-sparkmagic-conf` para personalizar o Sparkmagic, que atualiza valores no arquivo `config.json` para o Sparkmagic. Para obter mais informações sobre as configurações disponíveis, consulte example_config.json em. GitHub Para obter mais informações sobre como usar as classificações de configuração com aplicações no Amazon EMR, consulte Configurar aplicações.

O exemplo a seguir inicia um cluster usando o AWS CLI, referenciando o arquivo `MyJupyterConfig.json` para as configurações de classificação do Sparkmagic.

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr create-cluster --use-default-roles --release-label emr-5.14.0 \
--applications Name=Jupyter --instance-type m4.xlarge --instance-count 3 \
--ec2-attributes KeyName=MyKey,SubnetId=subnet-1234a5b6 --configurations file://
MyJupyterConfig.json
```

Os conteúdos de exemplo de `MyJupyterConfig.json` são os seguintes:

```
[
    {
    "Classification":"jupyter-sparkmagic-conf",
    "Properties": {
      "kernel_python_credentials" : "{\"username\":\"diego\",\"base64_password\":
\"mypass\",\"url\":\"http:\/\/localhost:8998\",\"auth\":\"None\"}"
      }
    }
]
```

> **Note**
>
> Com as versões 5.21.0 e posteriores do Amazon EMR, você pode substituir as configurações de cluster e especificar classificações de configuração adicionais para cada grupo de instâncias em um cluster em execução. Você faz isso usando o console do Amazon EMR, o AWS Command Line Interface (AWS CLI) ou o AWS SDK. Para obter mais informações, consulte Supplying a Configuration for an Instance Group in a Running Cluster.

Configurar a persistência de cadernos no Amazon S3

Você pode configurar um JupyterHub cluster no Amazon EMR para que os cadernos salvos por um usuário sejam mantidos no Amazon S3, fora do armazenamento temporário em instâncias do cluster. EC2

Especifique a persistência do Amazon S3 usando a classificação de configuração `jupyter-s3-conf` ao criar um cluster. Para obter mais informações, consulte Configurar aplicações.

Além de habilitar a persistência do Amazon S3 usando a propriedade `s3.persistence.enabled`, especifique um bucket no Amazon S3 no qual os cadernos são salvos com a propriedade `s3.persistence.bucket`. Os notebooks para cada usuário são salvos em uma pasta `jupyter/`*`jupyterhub-user-name`* no bucket especificado. O bucket já deve existir no Amazon S3 e a função para o perfil de EC2 instância que você especifica ao criar o cluster deve ter permissões para o bucket (por padrão, a função é`EMR_EC2_DefaultRole`). Para obter mais informações, consulte Configurar perfis do IAM para permissões do Amazon EMR para AWS serviços da.

Quando você inicia um novo cluster usando a mesma classificação de configuração propriedades, os usuários podem abrir notebooks com o conteúdo do local salvo.

Observe que, ao importar arquivos como módulos em um caderno quando o Amazon S3 está habilitado, isso resulta no carregamento dos arquivos para o Amazon S3. Quando você importa arquivos sem habilitar a persistência do Amazon S3, eles são carregados no contêiner. JupyterHub

O exemplo a seguir habilita a persistência do Amazon S3. Os notebooks salvos por usuários são salvos na pasta `s3://MyJupyterBackups/jupyter/`*`jupyterhub-user-name`* para cada usuário, em que *`jupyterhub-user-name`* é um nome de usuário, como `diego`.

```
[
    {
        "Classification": "jupyter-s3-conf",
        "Properties": {
            "s3.persistence.enabled": "true",
            "s3.persistence.bucket": "MyJupyterBackups"
        }
    }
]
```

Conectar-se ao nó principal e aos servidores de cadernos

JupyterHub administradores e usuários de notebooks devem se conectar ao nó principal do cluster usando um túnel SSH e, em seguida, conectar-se às interfaces web servidas pelo JupyterHub nó principal. Para obter mais informações sobre como configurar um túnel SSH e usar o túnel para substituir conexões da Web, consulte Conectar-se ao cluster no Guia de gerenciamento do Amazon EMR.

Por padrão, JupyterHub no Amazon, o EMR está disponível por meio da porta 9443 no nó principal. O JupyterHub proxy interno também atende às instâncias do notebook por meio da porta 9443. JupyterHub e as interfaces web do Jupyter podem ser acessadas usando uma URL com o seguinte padrão:

https: //:9443 ***`MasterNodeDNS`***

Você pode especificar uma porta diferente usando a propriedade `c.JupyterHub.port` no arquivo `jupyterhub_config.py`. Para obter mais informações, consulte Noções básicas sobre redes na JupyterHub documentação.

Por padrão, JupyterHub no Amazon EMR usa um certificado autoassinado para criptografia SSL usando HTTPS. Os usuários são solicitados a confiar no certificado autoassinado quando eles se conectarem. Você pode usar um certificado confiável e suas próprias chaves. Substitua o arquivo

de certificado padrão, `server.crt`, e arquivo chave `server.key` no diretório `/etc/jupyter/conf/` no nó principal com o seu próprio certificado e arquivos de chave. Use as propriedades `c.JupyterHub.ssl_key` e `c.JupyterHub.ssl_cert` no arquivo `jupyterhub_config.py` para especificar seus materiais SSL. Para obter mais informações, consulte Configurações de segurança na JupyterHub documentação. Depois de atualizar o `jupyterhub_config.py`, reinicie o contêiner.

JupyterHub configuração e administração

JupyterHub e componentes relacionados são executados dentro de um contêiner Docker chamado `jupyterhub` que executa o sistema operacional Ubuntu. Há várias maneiras para você administrar os componentes que são executados dentro do contêiner.

> **Warning**
>
> As personalizações que você executa no contêiner podem não persistir se o contêiner for reiniciado. Recomendamos que você faça script ou automatize a configuração de contêiner para que possa reproduzir personalizações com mais facilidade.

Administração usando a linha de comando

Quando conectado ao nó principal usando SSH, você pode emitir comandos usando a interface de linha de comando (CLI) do Docker e especificando o contêiner por nome (`jupyterhub`) ou ID. Por exemplo, o `sudo docker exec jupyterhub` *`command`* executa comandos reconhecidos pelo sistema operacional ou por um aplicativo em execução dentro do contêiner. Você pode usar esse método para adicionar usuários ao sistema operacional e instalar aplicativos e bibliotecas adicionais no contêiner Docker. Por exemplo, a imagem do contêiner padrão inclui Conda para instalação de pacote, portanto, você pode executar o seguinte comando na linha de comando do nó principal para instalar um aplicativo, Keras, dentro do contêiner:

```
sudo docker exec jupyterhub conda install keras
```

Administração por envio de etapas

As etapas são uma maneira de enviar trabalhos a um cluster. Você pode enviar etapas quando o cluster é iniciado ou pode enviá-las para um cluster em execução. Os comandos que você executa

na linha de comando podem ser enviados como etapas usando o `command-runner.jar`. Para obter mais informações, consulte Trabalhar com etapas usando a CLI e o console no Guia de gerenciamento do Amazon EMR e no Executar comandos e scripts em um cluster do Amazon EMR.

Por exemplo, você pode usar o AWS CLI comando a seguir em um computador local para instalar o Keras da mesma forma que fez na linha de comando do nó principal no exemplo anterior:

```
aws emr add-steps --cluster-id MyClusterID --steps Name="Command
 Runner",Jar="command-runner.jar",Args="/usr/bin/sudo","/usr/bin/
docker","exec","jupyterhub","conda","install","keras"
```

Além disso, você pode fazer o script de uma sequência de etapas, carregar o script no Amazon S3 e, em seguida usar `script-runner.jar` para executar o script quando você cria o cluster ou adiciona o script como uma etapa. Para obter mais informações, consulte Executar comandos e scripts em um cluster do Amazon EMR. Para obter um exemplo, consulte the section called "Exemplo: script Bash para adicionar vários usuários".

Administração usando REST APIs

Jupyter, JupyterHub, e o proxy HTTP para JupyterHub fornecer REST APIs que você pode usar para enviar solicitações. Para enviar solicitações para JupyterHub, você deve passar um token de API com a solicitação. Você pode usar o comando `curl` na linha de comando do nó principal para executar comandos REST. Para obter mais informações, consulte os seguintes recursos:

- Usando JupyterHub a API REST na documentação do JupyterHub, que inclui instruções para gerar tokens de API
- API do servidor do Caderno Jupyter GitHub
- configurable-http-proxyem GitHub

O exemplo a seguir demonstra o uso da API REST JupyterHub para obter uma lista de usuários. O comando passa um token administrativo gerado anteriormente e usa a porta padrão, 9443 JupyterHub, para canalizar a saída para jq para facilitar a visualização:

```
curl -XGET -s -k https://$HOST:9443/hub/api/users \
-H "Authorization: token $admin_token" | jq .
```

Adicionar usuários e administradores do Caderno Jupyter

Você pode usar um dos dois métodos para os usuários se autenticarem para que JupyterHub possam criar cadernos e, opcionalmente, administrar. JupyterHub O método mais fácil é usar o módulo JupyterHub de autenticação conectável (PAM). Além disso, JupyterHub no Amazon EMR é compatível com o Plugin do autenticador LDAP para LDAP, JupyterHub para obter identidades de usuários em um servidor LDAP, como um servidor do Microsoft Active Directory. Instruções e exemplos para adicionar usuários com cada método de autenticação são fornecidos nesta seção.

JupyterHub no Amazon EMR tem um usuário padrão com permissões de administrador. O nome de usuário é `jovyan` e a senha é `jupyter`. É altamente recomendável que você substitua o usuário por outro usuário que tenha permissões administrativas. Você pode fazer isso usando uma etapa ao criar o cluster ou se conectando ao nó principal quando o cluster estiver em execução.

Tópicos

- Usar autenticação PAM
- Usar autenticação LDAP
- Personificação do usuário

Usar autenticação PAM

A criação de usuários do PAM JupyterHub no Amazon EMR é um processo de duas etapas. A primeira etapa é adicionar usuários ao sistema operacional em execução no contêiner `jupyterhub` no nó principal e adicionar um diretório inicial do usuário correspondente para cada usuário. A segunda etapa é adicionar esses usuários do sistema operacional como JupyterHub usuários, um processo conhecido como whitelisting in. JupyterHub Depois que um JupyterHub usuário é adicionado, ele pode se conectar ao JupyterHub URL e fornecer suas credenciais do sistema operacional para acesso.

Quando um usuário faz login, JupyterHub abre a instância do servidor do notebook para esse usuário, que é salva no diretório inicial do usuário no nó principal, que é `/var/lib/jupyter/home/`*`username`*. Se não existir uma instância do servidor de cadernos, JupyterHub gerará uma instância do caderno diretório inicial do usuário. As seções a seguir demonstram como adicionar usuários individualmente ao sistema operacional e ao sistema operacional JupyterHub, seguido por um script bash rudimentar que adiciona vários usuários.

Adicionar um usuário do sistema operacional ao contêiner

O exemplo a seguir usa primeiro o comando useradd dentro do contêiner para adicionar um único usuário, diego e criar um diretório inicial para esse usuário. O segundo comando usa chpasswd para estabelecer uma senha de diego para este usuário. Os comandos são executados na linha de comando do nó principal enquanto conectado usando SSH. Você também pode executar esses comandos usando uma etapa conforme descrito anteriormente em Administração por envio de etapas.

```
sudo docker exec jupyterhub useradd -m -s /bin/bash -N diego
sudo docker exec jupyterhub bash -c "echo diego:diego | chpasswd"
```

Adicionando um JupyterHub usuário

Você pode usar o painel do administrador JupyterHub ou a API REST para adicionar usuários e administradores, ou apenas usuários.

Adicionar usuários e administradores usando o painel do administrador no JupyterHub

1. Conecte-se ao nó principal usando SSH e faça login em https: //:9443 *MasterNodeDNS* com uma identidade que tenha permissões de administrador.
2. Escolha Control Panel (Painel de Controle), Admin.
3. Escolha User (Usuário), Add Users (Adicionar usuários) ou escolha Admin, Add Admins (Adicionar admins).

Para adicionar um usuário usando a API REST

1. Conecte-se ao nó principal usando o SSH e use o seguinte comando no nó principal, ou execute o comando como uma etapa.
2. Adquira um token administrativo para fazer solicitações de API e substitua-o *AdminToken* na etapa seguinte por esse token.
3. Use o comando a seguir, *UserName* substituindo-o por um usuário do sistema operacional que tenha sido criado dentro do contêiner.

   ```
   curl -XPOST -H "Authorization: token AdminToken" "https://$(hostname):9443/hub/api/users/UserName
   ```

> **Note**
>
> Você é automaticamente adicionado como um usuário JupyterHub não administrador ao fazer login na interface da JupyterHub Web pela primeira vez.

Exemplo: script Bash para adicionar vários usuários

O exemplo de script bash a seguir reúne as etapas anteriores desta seção para criar vários JupyterHub usuários. O script pode ser executado diretamente no nó principal, ou pode ser carregado no Amazon S3 e, em seguida, executado como uma etapa.

O script primeiro estabelece uma matriz de nomes de usuário e usa o comando `jupyterhub token` para criar um token de API para o administrador padrão, jovyan. Em seguida, ele cria um usuário do sistema operacional no contêiner `jupyterhub` para cada usuário, atribuindo uma senha inicial a cada um que é igual ao seu nome de usuário. Por fim, ele chama a operação da API REST para criar cada usuário em JupyterHub. Ele transmite o token gerado anteriormente no script e envia a resposta REST para o `jq` para facilitar a visualização.

```
# Bulk add users to container and JupyterHub with temp password of username
set -x
USERS=(shirley diego ana richard li john mary anaya)
TOKEN=$(sudo docker exec jupyterhub /opt/conda/bin/jupyterhub token jovyan | tail -1)
for i in "${USERS[@]}";
do
   sudo docker exec jupyterhub useradd -m -s /bin/bash -N $i
   sudo docker exec jupyterhub bash -c "echo $i:$i | chpasswd"
   curl -XPOST --silent -k https://$(hostname):9443/hub/api/users/$i \
 -H "Authorization: token $TOKEN" | jq
done
```

Salve o script em um local do Amazon S3, como `s3://amzn-s3-demo-bucket/createjupyterusers.sh`. Em seguida, você pode usar o `script-runner.jar` para executá-lo como uma etapa.

Exemplo: executar o script ao criar um cluster (AWS CLI)

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr create-cluster --name="MyJupyterHubCluster" --release-label emr-5.36.2 \
--applications Name=JupyterHub --log-uri s3://amzn-s3-demo-bucket/MyJupyterClusterLogs
 \
--use-default-roles --instance-type m5.xlarge --instance-count 2 --ec2-attributes
 KeyName=MyKeyPair \
--steps Type=CUSTOM_JAR,Name=CustomJAR,ActionOnFailure=CONTINUE,\
Jar=s3://region.elasticmapreduce/libs/script-runner/script-runner.jar,Args=["s3://amzn-
s3-demo-bucket/createjupyterusers.sh"]
```

Executar o script em um cluster existente (AWS CLI)

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr add-steps --cluster-id j-XXXXXXXX --steps Type=CUSTOM_JAR,\
Name=CustomJAR,ActionOnFailure=CONTINUE,\
Jar=s3://region.elasticmapreduce/libs/script-runner/script-runner.jar,Args=["s3://amzn-
s3-demo-bucket/createjupyterusers.sh"]
```

Usar autenticação LDAP

O LDAP (Lightweight Directory Access Protocol) é um protocolo de aplicativo para consultar e modificar objetos que correspondem a recursos, como usuários e computadores, armazenados em um provedor de serviços de diretório compatível com LDAP, como o Active Directory ou um servidor OpenLDAP. Você pode usar o Plugin do autenticador LDAP JupyterHub no JupyterHub Amazon

EMR para usar LDAP para autenticação do usuário. O plug-in processa sessões de login para usuários LDAP e fornece as informações do usuário para o Jupyter. Isso permite que os usuários se conectem a notebooks JupyterHub e notebooks usando as credenciais de suas identidades armazenadas em um servidor compatível com LDAP.

As etapas desta seção orientam você nas etapas a seguir para configurar e habilitar o LDAP usando o Plugin do autenticador LDAP para. JupyterHub Execute as etapas enquanto estiver conectado à linha de comando do nó principal. Para obter mais informações, consulte Conectar-se ao nó principal e aos servidores de cadernos.

1. Crie um arquivo de configuração LDAP com informações sobre o servidor LDAP, como o endereço IP do host, a porta, nomes de vinculação, e assim por diante.
2. Modifique o `/etc/jupyter/conf/jupyterhub_config.py` para habilitar o Plugin do Autenticador LDAP para o JupyterHub.
3. Crie e execute um script que configura o LDAP dentro do contêiner `jupyterhub`.
4. Consulte o LDAP para usuários e, em seguida, crie diretórios pessoais dentro do contêiner para cada usuário. JupyterHub requer diretórios pessoais para hospedar notebooks.
5. Execute um script que reinicia JupyterHub

> **Important**
>
> Antes de configurar o LDAP, teste a infraestrutura de rede para garantir que o servidor LDAP e o nó principal do cluster possa se comunicar, conforme necessário. O TLS normalmente usa a porta 389 por meio de uma conexão TCP simples. Se a sua conexão LDAP usar SSL, a porta TCP conhecida para SSL será 636.

Criar o arquivo de configuração do LDAP

O exemplo abaixo usa os seguintes valores de configuração do espaço reservado. Substitua esses valores pelos parâmetros que corresponde à sua implementação.

- O servidor LDAP está executando a versão 3 e está disponível na porta 389. Essa é a porta não SSL padrão para o LDAP.
- O nome diferenciado base (DN) é `dc=example, dc=org`.

Use um editor de texto para criar o arquivo ldap.conf, com conteúdo semelhante ao seguinte. Use os valores apropriados para a sua implementação do LDAP. *host*Substitua pelo endereço IP ou nome de host resolvível do seu servidor LDAP.

```
base dc=example,dc=org
uri ldap://host
ldap_version 3
binddn cn=admin,dc=example,dc=org
bindpw admin
```

Habilitar o Plugin do autenticador LDAP para JupyterHub

Use um editor de texto para modificar o arquivo `/etc/jupyter/conf/jupyterhub_config.py` e adicionar as propriedades ldapauthenticator semelhante ao seguinte. *host*Substitua pelo endereço IP ou o nome do host que pode ser resolvido do servidor LDAP. O exemplo pressupõe que os objetos do usuário estão dentro de uma unidade organizacional (ou) chamada *people* e usa os componentes de nome distinto que você estabeleceu anteriormente usando`ldap.conf`.

```
c.JupyterHub.authenticator_class = 'ldapauthenticator.LDAPAuthenticator'
c.LDAPAuthenticator.use_ssl = False
c.LDAPAuthenticator.server_address = 'host'
c.LDAPAuthenticator.bind_dn_template = 'cn={username},ou=people,dc=example,dc=org'
```

Configurar o LDAP dentro do contêiner

Use um editor de texto para criar um script bash com o seguinte conteúdo:

```
#!/bin/bash

# Uncomment the following lines to install LDAP client libraries only if
# using Amazon EMR release version 5.14.0. Later versions install libraries by default.
# sudo docker exec jupyterhub bash -c "sudo apt-get update"
# sudo docker exec jupyterhub bash -c "sudo apt-get -y install libnss-ldap libpam-ldap
 ldap-utils nscd"

# Copy ldap.conf
sudo docker cp ldap.conf jupyterhub:/etc/ldap/
sudo docker exec jupyterhub bash -c "cat /etc/ldap/ldap.conf"

# configure nss switch
sudo docker exec jupyterhub bash -c "sed -i 's/\(^passwd.*\)/\1 ldap/g' /etc/
nsswitch.conf"
```

```
sudo docker exec jupyterhub bash -c "sed -i 's/\(^group.*\)/\1 ldap/g' /etc/
nsswitch.conf"
sudo docker exec jupyterhub bash -c "sed -i 's/\(^shadow.*\)/\1 ldap/g' /etc/
nsswitch.conf"
sudo docker exec jupyterhub bash -c "cat /etc/nsswitch.conf"

# configure PAM to create home directories
sudo docker exec jupyterhub bash -c "echo 'session required        pam_mkhomedir.so
 skel=/etc/skel umask=077' >> /etc/pam.d/common-session"
sudo docker exec jupyterhub bash -c "cat /etc/pam.d/common-session"

# restart nscd service
sudo docker exec jupyterhub bash -c "sudo service nscd restart"

# Test
sudo docker exec jupyterhub bash -c "getent passwd"

# Install ldap plugin
sudo docker exec jupyterhub bash -c "pip install jupyterhub-ldapauthenticator"
```

Salve o script no nó principal e, em seguida, execute-o na linha de comando do nó principal. Por exemplo, com o script salvo como `configure_ldap_client.sh`, torne o arquivo executável:

```
chmod +x configure_ldap_client.sh
```

E execute o script:

```
./configure_ldap_client.sh
```

Adicionar atributos ao Active Directory

Para encontrar cada usuário e criar a entrada apropriada no banco de dados, o contêiner JupyterHub docker requer as seguintes propriedades UNIX para o objeto de usuário correspondente no Active Directory. Para obter mais informações, consulte a seção Como continuo a editar os atributos GID/UID RFC 2307 agora que o Plug-in de atributos do Unix não está mais disponível para o snap-in do MMC de usuários e computadores do Active Directory? no artigo Esclarecimento sobre o status do gerenciamento de identidade para Unix (IDMU) da função do servidor NIS na visualização técnica do Windows Server 2016 e além.

- `homeDirectory`

Este é o local para o diretório inicial do usuário, que é geralmente `/home/`*`username`*.

- `gidNumber`

  Um valor maior que 60000 que ainda não esteja sendo usado por outro usuário. Verifique o arquivo `etc/passwd` para ver os GIDs que estão sendo usados.

- `uidNumber`

  Um valor maior que 60000 que ainda não esteja sendo usado por outro grupo. Verifique o arquivo `etc/group` para ver os UDIs que estão sendo usados.

- `uid`

  Isso é o mesmo que *`username`* o.

Criar diretórios iniciais do usuário

JupyterHub precisa de diretórios iniciais dentro do contêiner para autenticar usuários do LDAP e armazenar dados da instância. O exemplo a seguir demonstra dois usuários, shirley e diego, no diretório LDAP.

A primeira etapa é consultar o servidor LDAP para obter as informações de ID de usuário e ID de grupo de cada usuário usando ldapsearch, conforme mostrado no exemplo a seguir, *`host`* substituindo pelo endereço IP ou nome de host resolvível do seu servidor LDAP:

```
ldapsearch -x -H ldap://host \
 -D "cn=admin,dc=example,dc=org" \
 -w admin \
 -b "ou=people,dc=example,dc=org" \
 -s sub \
 "(objectclass=*)" uidNumber gidNumber
```

O comando `ldapsearch` retorna uma resposta formatada LDIF que se parece com a seguinte para os usuários shirley e diego.

```
# extended LDIF

# LDAPv3
# base <ou=people,dc=example,dc=org> with scope subtree
# filter: (objectclass=*)
```

```
# requesting: uidNumber gidNumber sn

# people, example.org
dn: ou=people,dc=example,dc=org

# diego, people, example.org
dn: cn=diego,ou=people,dc=example,dc=org
sn: B
uidNumber: 1001
gidNumber: 100

# shirley, people, example.org
dn: cn=shirley,ou=people,dc=example,dc=org
sn: A
uidNumber: 1002
gidNumber: 100

# search result
search: 2
result: 0 Success

# numResponses: 4
# numEntries: 3
```

Ao usar as informações da resposta, execute comandos dentro do contêiner para criar um diretório inicial para cada nome comum de usuário (`cn`). Use o `uidNumber` e o `gidNumber` para corrigir a propriedade do diretório inicial desse usuário. Os comandos de exemplo a seguir fazem isso para o usuário*shirley*.

```
sudo docker container exec jupyterhub bash -c "mkdir /home/shirley"
sudo docker container exec jupyterhub bash -c "chown -R $uidNumber /home/shirley"
sudo docker container exec jupyterhub bash -c "sudo chgrp -R $gidNumber /home/shirley"
```

> **Note**
>
> O autenticador LDAP para JupyterHub não oferece suporte à criação de usuários locais. Para obter mais informações, consulte Observação da configuração do autenticador do LDAP sobre criação de usuário local.
> Para criar um usuário local manualmente, use o comando a seguir.

```
sudo docker exec jupyterhub bash -c "echo 'shirley:x:$uidNumber:$gidNumber::/
home/shirley:/bin/bash' >> /etc/passwd"
```

Reiniciar o JupyterHub contêiner

Para reiniciar o contêiner `jupyterhub`, execute os seguintes comandos:

```
sudo docker stop jupyterhub
sudo docker start jupyterhub
```

Personificação do usuário

Um trabalho do Spark em execução em um caderno Jupyter percorre várias aplicações durante sua execução no Amazon EMR. Por exemplo, PySpark 3 códigos que um usuário executa dentro do Jupyter são recebidos pelo Sparkmagic, que usa uma solicitação HTTP POST para enviá-los ao Livy, que então cria um trabalho do Spark para executar no cluster usando o YARN.

Por padrão, os trabalhos do YARN que são enviados dessa forma são executados como usuário `livy`, independentemente do usuário que iniciou o trabalho. Ao configurar a personificação do usuário, você pode fazer com que o ID de usuário do notebook também seja o usuário associado ao trabalho do YARN. Em vez dos trabalhos serem iniciados por `shirley` e `diego` associados ao usuário `livy`, os trabalhos que cada usuário inicia são associados a `shirley` e `diego` respectivamente. Isso ajuda você a auditar o uso do Jupyter e gerenciar os aplicativos em sua organização.

Essa configuração tem suporte somente quando as chamadas do Sparkmagic para o Livy são não autenticadas. Os aplicativos que fornecem uma camada de proxy ou de autenticação entre os aplicativos Hadoop e Livy (como o Apache Knox Gateway) não são compatíveis. As etapas para configurar a representação do usuário nesta seção pressupõem que JupyterHub e Livy estejam sendo executados no mesmo nó principal. Se o seu aplicativo tiver clusters separados, Etapa 3: criar diretórios iniciais do HDFS para usuários precisará ser modificado para que os diretórios do HDFS sejam criados no nó principal do Livy.

Etapas para configurar a personificação do usuário

- Etapa 1: configurar o Livy
- Etapa 2: adicionar usuários

- Etapa 3: criar diretórios iniciais do HDFS para usuários

Etapa 1: configurar o Livy

Use as classificações de configuração `livy-conf` e `core-site` ao criar um cluster para habilitar a personificação do usuário do Livy, conforme mostrado no exemplo a seguir. Salve a classificação de configuração como um JSON e, em seguida, faça referência a ele quando criar o cluster ou especifique a classificação de configuração em linha. Para obter mais informações, consulte Configurar aplicações.

```
[
  {
    "Classification": "livy-conf",
    "Properties": {
      "livy.impersonation.enabled": "true"
    }
  },
  {
    "Classification": "core-site",
    "Properties": {
      "hadoop.proxyuser.livy.groups": "*",
      "hadoop.proxyuser.livy.hosts": "*"
    }
  }
]
```

Etapa 2: adicionar usuários

Adicione JupyterHub usuários usando PAM ou LDAP. Para obter mais informações, consulte Usar autenticação PAM e Usar autenticação LDAP.

Etapa 3: criar diretórios iniciais do HDFS para usuários

Você se conectou ao nó principal para criar usuários. Ainda conectado a ele, copie o conteúdo abaixo e o salve em um arquivo de script. O script cria diretórios iniciais do HDFS para cada JupyterHub usuário no nó principal. O script pressupõe que você esteja usando o ID de usuário administrador padrão,*jovyan*.

```
#!/bin/bash

CURL="curl --silent -k"
```

```
HOST=$(curl -s http://169.254.169.254/latest/meta-data/local-hostname)

admin_token() {
    local user=jovyan
    local pwd=jupyter
    local token=$($CURL https://$HOST:9443/hub/api/authorizations/token \
        -d "{\"username\":\"$user\", \"password\":\"$pwd\"}" | jq ".token")
    if [[ $token != null ]]; then
        token=$(echo $token | sed 's/"//g')
    else
        echo "Unable to get Jupyter API Token."
        exit 1
    fi
    echo $token
}

# Get Jupyter Admin token
token=$(admin_token)

# Get list of Jupyter users
users=$(curl -XGET -s -k https://$HOST:9443/hub/api/users \
 -H "Authorization: token $token" | jq '.[].name' | sed 's/"//g')

# Create HDFS home dir
for user in ${users[@]};
do
 echo "Create hdfs home dir for $user"
 hadoop fs -mkdir /user/$user
 hadoop fs -chmod 777 /user/$user
done
```

Instalar Kernels e bibliotecas adicionais

Quando você cria um cluster JupyterHub no Amazon EMR, o kernel padrão do Python 3 do Python 3 para Jupyter e os kernels do PySpark para Spark para Spark para Spark são PySpark instalados no contêiner do Docker. Você pode instalar kernels adicionais. Você também pode instalar bibliotecas e pacotes adicionais e importá-los para o shell apropriado.

Instalar um Kernel

Os kernels são instalados no contêiner Docker. A maneira mais fácil de fazer isso é criar um script bash com comandos de instalação, salvá-lo no nó principal e, em seguida, usar o comando

`sudo docker exec jupyterhub` *`script_name`* para executar o script dentro do contêiner `jupyterhub`. O script de exemplo a seguir instala o kernel e, em seguida, instala algumas bibliotecas para esse kernel no nó principal, para que mais tarde você possa importar as bibliotecas usando o kernel no Jupyter.

```
#!/bin/bash

# Install Python 2 kernel
conda create -n py27 python=2.7 anaconda
source /opt/conda/envs/py27/bin/activate
apt-get update
apt-get install -y gcc
/opt/conda/envs/py27/bin/python -m pip install --upgrade ipykernel
/opt/conda/envs/py27/bin/python -m ipykernel install

# Install libraries for Python 2
/opt/conda/envs/py27/bin/pip install paramiko nltk scipy numpy scikit-learn pandas
```

Para instalar o kernel e as bibliotecas dentro do contêiner, abra uma conexão de terminal com o nó principal, salve o script em `/etc/jupyter/install_kernels.sh` e execute o seguinte comando na linha de comando do nó principal:

```
sudo docker exec jupyterhub bash /etc/jupyter/install_kernels.sh
```

Usar bibliotecas e instalar bibliotecas adicionais

Um conjunto principal de machine learning e bibliotecas de ciência de dados para Python 3 são pré-instalados no JupyterHub Amazon EMR. Você pode usar o `sudo docker exec jupyterhub bash -c "conda list"` e o `sudo docker exec jupyterhub bash -c "pip freeze"`.

Se um job Spark precisar de bibliotecas nos nós de operador, recomendamos que você use uma ação de bootstrap para executar um script para instalar as bibliotecas ao criar o cluster. As ações de bootstrap são executadas em todos os nós do cluster durante o processo de criação do cluster, o que simplifica a instalação. Se você instalar bibliotecas em nós core/de operador depois que um cluster estiver em execução, a operação será mais complexa. Nós fornecemos um exemplo de programa em Python nesta seção que mostra como instalar essas bibliotecas.

A ação de bootstrap e os exemplos de programas em Python mostrados nesta seção usam um script bash salvo no Amazon S3 para instalar as bibliotecas em todos os nós.

O script referenciado nos exemplos a seguir usa `pip` para instalar paramiko, nltk, scipy, scikit-learn e pandas para o kernel do Python 3:

```
#!/bin/bash

sudo python3 -m pip install boto3 paramiko nltk scipy scikit-learn pandas
```

Depois de criar o script, carregue-o em um local no Amazon S3, por exemplo, `s3://amzn-s3-demo-bucket/install-my-jupyter-libraries.sh`. Para obter mais informações, consulte Carregar objetos no Guia do usuário do Amazon Simple Storage Service para que você possa usá-lo na sua ação de bootstrap ou no seu programa do Python.

Especificar uma ação de bootstrap que instala bibliotecas em todos os nós ao criar um cluster usando a AWS CLI

1. Crie um script semelhante ao do exemplo anterior e salve-o em um local no Amazon S3. Nós usamos o exemplo `s3://amzn-s3-demo-bucket/install-my-jupyter-libraries.sh`.
2. Crie o cluster com JupyterHub o `Path` argumento da `--bootstrap-actions` opção para especificar a localização do script, conforme mostrado no exemplo a seguir:

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr create-cluster --name="MyJupyterHubCluster" --release-label emr-5.36.2 \
--applications Name=JupyterHub --log-uri s3://amzn-s3-demo-bucket/
MyJupyterClusterLogs \
--use-default-roles --instance-type m5.xlarge --instance-count 2 --ec2-attributes
 KeyName=MyKeyPair \
--bootstrap-actions Path=s3://amzn-s3-demo-bucket/install-my-jupyter-
libraries.sh,Name=InstallJupyterLibs
```

Como especificar uma ação de bootstrap que instala bibliotecas em todos os nós ao criar um cluster usando o console

1. Navegue até o novo console do Amazon EMR e selecione Alternar para o console antigo na navegação lateral. Para obter mais informações sobre o que esperar ao alternar para o console antigo, consulte Usar o console antigo.
2. Escolha Create cluster (Criar cluster), Go to advanced options (Ir para opções avançadas).
3. Especifique as configurações para Software and Steps (Software e etapas) e Hardware conforme apropriado para seu aplicativo.
4. Na tela General Cluster Settings (Configurações gerais de cluster), expanda Bootstrap Actions (Ações de bootstrap).
5. Para Add bootstrap action (Adicionar ação de bootstrap), selecione Custom action (Ação personalizada), Configure and add (Configurar e adicionar).
6. Em Nome, insira um nome fácil de lembrar. Em Local do script, insira o local no Amazon S3 do seu script (o exemplo que usamos é s3://amzn-s3-demo-bucket/ install-my-jupyter-libraries .sh). Deixe Optional arguments (Argumentos opcionais) em branco e escolha Add (Adicionar).
7. Especifique outras configurações para o seu cluster e escolha Next (Próximo).
8. Especifique as configurações de segurança e escolha Create cluster (Criar cluster).

Example Instalar bibliotecas em nós centrais de um cluster em execução

Após instalar bibliotecas no nó principal de dentro do Jupyter, você pode instalar bibliotecas em nós core em execução de várias maneiras. O exemplo a seguir mostra um programa Python escrito para ser executado em uma máquina local. Quando você executa o programa Python localmente, ele usa o `AWS-RunShellScript` of AWS Systems Manager para executar o script de exemplo, mostrado anteriormente nesta seção, que instala bibliotecas nos nós principais do cluster.

```
import argparse
import time
import boto3


def install_libraries_on_core_nodes(cluster_id, script_path, emr_client, ssm_client):
    """
    Copies and runs a shell script on the core nodes in the cluster.

    :param cluster_id: The ID of the cluster.
```

```
    :param script_path: The path to the script, typically an Amazon S3 object URL.
    :param emr_client: The Boto3 Amazon EMR client.
    :param ssm_client: The Boto3 AWS Systems Manager client.
    """
    core_nodes = emr_client.list_instances(
        ClusterId=cluster_id, InstanceGroupTypes=["CORE"]
    )["Instances"]
    core_instance_ids = [node["Ec2InstanceId"] for node in core_nodes]
    print(f"Found core instances: {core_instance_ids}.")

    commands = [
        # Copy the shell script from Amazon S3 to each node instance.
        f"aws s3 cp {script_path} /home/hadoop",
        # Run the shell script to install libraries on each node instance.
        "bash /home/hadoop/install_libraries.sh",
    ]
    for command in commands:
        print(f"Sending '{command}' to core instances...")
        command_id = ssm_client.send_command(
            InstanceIds=core_instance_ids,
            DocumentName="AWS-RunShellScript",
            Parameters={"commands": [command]},
            TimeoutSeconds=3600,
        )["Command"]["CommandId"]
        while True:
            # Verify the previous step succeeded before running the next step.
            cmd_result = ssm_client.list_commands(CommandId=command_id)["Commands"][0]
            if cmd_result["StatusDetails"] == "Success":
                print(f"Command succeeded.")
                break
            elif cmd_result["StatusDetails"] in ["Pending", "InProgress"]:
                print(f"Command status is {cmd_result['StatusDetails']}, waiting...")
                time.sleep(10)
            else:
                print(f"Command status is {cmd_result['StatusDetails']}, quitting.")
                raise RuntimeError(
                    f"Command {command} failed to run. "
                    f"Details: {cmd_result['StatusDetails']}"
                )


def main():
    parser = argparse.ArgumentParser()
    parser.add_argument("cluster_id", help="The ID of the cluster.")
```

```
    parser.add_argument("script_path", help="The path to the script in Amazon S3.")
    args = parser.parse_args()

    emr_client = boto3.client("emr")
    ssm_client = boto3.client("ssm")

    install_libraries_on_core_nodes(
        args.cluster_id, args.script_path, emr_client, ssm_client
    )


if __name__ == "__main__":
    main()
```

JupyterHub histórico de versões

A tabela a seguir lista a versão do JupyterHub incluída em cada versão do Amazon EMR, junto com os componentes instalados com a aplicação. Para obter as versões dos componentes em cada versão, consulte a seção Versão de componentes da sua versão em Versões 7.x do Amazon EMR, Versões de lançamento 6.x do Amazon EMR ou Versões de lançamento 5.x do Amazon EMR.

JupyterHub informações sobre a versão

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-7.9.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-7.8.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-7.7.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-7.6.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-7.5.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-7.4.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-7.3.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-7.2.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.36.2	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-7.1.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-7.0.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-6.15.0	1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.14.0	1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-6.13.0	1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.12.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-6.11.1	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.11.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-6.10.1	1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.10.0	1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-6.9.1	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.9.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-6.8.1	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.8.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-6.7.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.36.1	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-5.36.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.6.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-5.35.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.5.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-6.4.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.3.1	1.2.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-6.3.0	1.2.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.2.1	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-6.2.0	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.1.1	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-6.1.0	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.0.1	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-6.0.0	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.34.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-5.33.1	1.2.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.33.0	1.2.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-5.32.1	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.32.0	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-5.31.1	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.31.0	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-5.30.2	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.30.1	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-5.30.0	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.29.0	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-5.28.1	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.28.0	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-5.27.1	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.27.0	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-5.26.0	0.9.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.25.0	0.9.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-5.24.1	0.9.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.24.0	0.9.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-5.23.1	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.23.0	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-5.22.0	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.21.2	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-5.21.1	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.21.0	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-5.20.1	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.20.0	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-5.19.1	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.19.0	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-5.18.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.18.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-5.17.2	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.17.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-5.17.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.16.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-5.16.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.15.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-5.15.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.14.2	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Rótulo de versão do Amazon EMR	JupyterHub Versão	Componentes instalados com JupyterHub
emr-5.14.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.14.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Apache Livy

O Livy permite a interação por meio de uma interface REST com um cluster do EMR executando o Spark. Você pode usar a interface REST ou uma biblioteca de cliente RPC para enviar trabalhos do Spark ou fragmentos de código do Spark, recuperar os resultados de forma síncrona e assíncrona e gerenciar o Spark Context. Para obter mais informações, consulte o site do Apache Livy. O Livy é incluído nas versões 5.9.0 e posteriores do Amazon EMR.

Para acessar a interface da web Livy, configure um túnel SSH para o nó principal e uma conexão de proxy. Para obter mais informações, consulte Visualizar interfaces da Web hospedadas em clusters do EMR.

A tabela a seguir lista a versão do Livy incluída na versão mais recente da série 7.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Livy.

Para obter a versão dos componentes instalados com o Livy nesse lançamento, consulte Release 7.9.0 Component Versions.

Informações de versão do Livy para o emr-7.9.0

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-7.9.0	Livy 0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

A tabela a seguir lista a versão do Livy incluída na versão mais recente da série 6.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Livy.

Para obter a versão dos componentes instalados com o Livy nessa versão, consulte Release 6.15.0 Component Versions.

Informações de versão do Livy para o emr-6.15.0

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-6.15.0	Livy 0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

A tabela a seguir lista a versão do Livy incluída na versão mais recente da série 5.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Livy.

Para obter a versão dos componentes instalados com o Livy nesse lançamento, consulte Release 5.36.2 Component Versions.

Informações de versão do Livy para o emr-5.36.2

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.36.2	Livy 0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
		yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Tópicos

- Habilitar HTTPS com o Apache Livy
- Histórico de versões do Livy

Habilitar HTTPS com o Apache Livy

Este tópico é relevante se você estiver executando o Amazon EMR 7.3.0 ou uma versão anterior. A partir da versão 7.4.0, o HTTPS é habilitado com o Apache Livy por padrão.

1. Provisione um cluster do Amazon EMR com criptografia em trânsito habilitada. Para saber mais sobre criptografia, consulte Criptografar dados em repouso e em trânsito.
2. Crie um arquivo chamado `livy_ssl.sh` com o conteúdo a seguir.

```
#!/bin/bash

KEYSTORE_FILE=`awk '/ssl.server.keystore.location/{getline; print}' /etc/hadoop/
conf/ssl-server.xml | sed -e 's/<[^>]*>//g' | tr -d ' \t\n\r\f'`
KEYSTORE_PASS=`awk '/ssl.server.keystore.password/{getline; print}' /etc/hadoop/
conf/ssl-server.xml | sed -e 's/<[^>]*>//g' | tr -d ' \t\n\r\f'`
KEY_PASS=`awk '/ssl.server.keystore.keypassword/{getline; print}' /etc/hadoop/conf/
ssl-server.xml | sed -e 's/<[^>]*>//g' | tr -d ' \t\n\r\f'`

echo "livy.keystore $KEYSTORE_FILE
livy.keystore.password $KEYSTORE_PASS
livy.key-password $KEY_PASS" | sudo tee -a /etc/livy/conf/livy.conf >/dev/null
```

```
sudo systemctl restart livy-server.service
```

3. Execute o script a seguir como uma etapa do Amazon EMR. Esse script modifica `/etc/livy/conf/livy.conf` para ativar o SSL.

```
--steps '[{"Args":["s3://amzn-s3-demo-bucket/
livy_ssl.sh"],"Type":"CUSTOM_JAR","ActionOnFailure":"CONTINUE","Jar":"s3://
us-east-1.elasticmapreduce/libs/script-runner/script-
runner.jar","Properties":"","Name":"Custom JAR"}]'
```

4. Reinicialize o serviço Apache Livy para que as alterações surtem efeito. Para reiniciar o Apache Livy, consulte Interromper e reiniciar processos.
5. Teste se os clientes agora podem se comunicar usando HTTPS. Para enviar um trabalho, por exemplo, execute o código a seguir.

```
curl -k -X POST --data '{"file": "local:///usr/lib/spark/examples/jars/spark-
examples.jar",
"className": "org.apache.spark.examples.SparkPi"}' \
-H "Content-Type: application/json" \
https://EMR_Master_Node_Host:8998/batches
```

Se você tiver habilitou o HTTPS com sucesso, o Livy enviará uma resposta indicando que o comando foi aceito e que o trabalho em lotes foi enviado.

```
{"id":1,"name":null,"owner":null,"proxyUser":null,"state":"starting","appId":null,"appInfo"
{"driverLogUrl":null,"sparkUiUrl":null},"log":["stdout: ","\nstderr: ","\nYARN
 Diagnostics: "]}
```

Histórico de versões do Livy

A tabela a seguir lista a versão do Livy incluída em cada versão do Amazon EMR, além dos componentes instalados com a aplicação. Para obter as versões dos componentes em cada versão, consulte a seção Versão de componentes da sua versão em Versões 7.x do Amazon EMR, Versões de lançamento 6.x do Amazon EMR ou Versões de lançamento 5.x do Amazon EMR.

Informações das versões do Livy

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-7.9.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-7.8.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-7.7.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager,

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
		hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-7.6.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-7.5.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-7.4.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-7.3.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-7.2.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.36.2	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-7.1.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-7.0.0	0.7.1	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-6.15.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.14.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-6.13.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.12.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-6.11.1	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.11.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-6.10.1	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.10.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-6.9.1	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.9.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-6.8.1	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.8.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-6.7.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.36.1	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.36.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.6.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.35.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.5.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-6.4.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.3.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-6.3.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.2.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-6.2.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.1.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-6.1.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.0.1	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-6.0.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.34.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.33.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.33.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.32.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.32.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.31.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.31.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.30.2	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.30.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.30.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-notebook-env, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.29.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.28.1	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.28.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.27.1	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.27.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.26.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.25.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.24.1	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.24.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.23.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.23.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.22.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.21.2	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.21.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.21.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.20.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.20.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.19.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.19.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.18.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.18.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.17.2	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.17.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.17.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.16.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.16.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.15.1	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.15.0	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.14.2	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.14.1	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.14.0	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.13.1	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.13.0	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.12.3	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.12.2	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.12.1	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.12.0	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.11.4	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.11.3	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.11.2	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.11.1	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.11.0	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.10.1	0.4.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.10.0	0.4.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.9.1	0.4.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Rótulo de versão do Amazon EMR	Versão do Livy	Componentes instalados com o Livy
emr-5.9.0	0.4.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Apache MXNet

> **Note**
>
> MXNet chegou ao fim da vida útil e o projeto está arquivado. O Amazon EMR 7.2.0 é a última versão que inclui. MXNet

O Apache MXNet é uma biblioteca de aceleração projetada para criar redes neurais e outros aplicativos de aprendizado profundo. MXNetautomatiza fluxos de trabalho comuns e otimiza os cálculos numéricos. MXNet ajuda você a projetar arquiteturas de redes neurais sem precisar se concentrar na implementação de cálculos de baixo nível, como operações de álgebra linear. MXNet está incluído na versão 5.10.0 e posterior do Amazon EMR.

Para obter mais informações, consulte o MXNet site da Apache.

A tabela a seguir lista a versão MXNet incluída na versão mais recente da série 7.x do Amazon EMR, junto com os componentes com os quais o Amazon EMR é instalado. MXNet

Para a versão dos componentes instalados com MXNet, consulte Versões de componentes da versão 7.2.0.

MXNet informações sobre a versão do emr-7.2.0

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-7.2.0	MXNet 1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

A tabela a seguir lista a versão MXNet incluída na versão mais recente da série 6.x do Amazon EMR, junto com os componentes com os quais o Amazon EMR é instalado. MXNet

Para a versão dos componentes instalados MXNet nesta versão, consulte Versões de componentes da versão 6.15.0.

MXNet informações da versão do emr-6.15.0

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-6.15.0	MXNet 1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

A tabela a seguir lista a versão MXNet incluída na versão mais recente da série 5.x do Amazon EMR, junto com os componentes com os quais o Amazon EMR é instalado. MXNet

Para a versão dos componentes instalados MXNet nesta versão, consulte Versões de componentes da versão 5.36.2.

MXNet informações da versão do emr-5.36.2

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-5.36.2	MXNet 1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resour

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
		cemanager, hadoop-yarn-timeline-server, mxnet, opencv

MXNet histórico de lançamentos

A tabela a seguir lista a versão MXNet incluída em cada versão de lançamento do Amazon EMR, junto com os componentes instalados com o aplicativo. Para obter as versões dos componentes em cada versão, consulte a seção Versão de componentes da sua versão em Versões 7.x do Amazon EMR, Versões de lançamento 6.x do Amazon EMR ou Versões de lançamento 5.x do Amazon EMR.

MXNet informações sobre a versão

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-7.2.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.36.2	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resour

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
		cemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-7.1.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-7.0.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.15.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-6.14.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.13.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.12.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-6.11.1	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.11.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.10.1	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-6.10.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.9.1	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.9.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-6.8.1	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.8.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.7.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-5.36.1	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.36.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.6.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-5.35.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.5.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.4.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-6.3.1	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.3.0	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.2.1	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-6.2.0	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.1.1	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.1.0	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-6.0.1	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.0.0	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.34.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-5.33.1	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.33.0	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.32.1	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-5.32.0	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.31.1	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.31.0	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-5.30.2	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.30.1	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.30.0	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-5.29.0	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.28.1	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.28.0	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-5.27.1	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.27.0	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.26.0	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-5.25.0	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.24.1	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.24.0	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-5.23.1	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.23.0	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.22.0	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-5.21.2	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.21.1	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.21.0	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-5.20.1	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.20.0	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.19.1	1.3.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-5.19.0	1.3.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.18.1	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.18.0	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-5.17.2	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.17.1	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.17.0	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-5.16.1	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.16.0	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.15.1	1.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-5.15.0	1.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.14.2	1.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.14.1	1.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-5.14.0	1.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.13.1	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.13.0	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-5.12.3	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.12.2	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.12.1	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-5.12.0	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.11.4	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.11.3	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-5.11.2	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.11.1	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.11.0	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet

Rótulo de versão do Amazon EMR	MXNet Versão	Componentes instalados com MXNet
emr-5.10.1	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.10.0	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet

Apache Oozie

> **Note**
>
> Suspensão de uso do Oozie — Como o OSS Oozie entrou no modo sótão, ele será removido de uma das próximas versões do EMR. Recomendamos que os clientes usem o MWAA como alternativa ao Oozie. Você pode consultar o utilitário para GitHub migrar os fluxos oozie-to-airflow-emrde trabalho existentes do Oozie para o MWAA.

Use o programador de fluxo de trabalho do Apache Oozie para gerenciar e coordenar trabalhos do Hadoop. Para obter mais informações, consulte https://attic.apache.org/projects/oozie.html.

A interface nativa da Web do Oozie não é compatível com o Amazon EMR. Para usar uma interface de front-end para o Oozie, tente o aplicativo Hue do Oozie. Para obter mais informações, consulte Hue. O Oozie é incluído com as versões 5.0.0 e posteriores do Amazon EMR. Oozie é incluído como um aplicativo em sandbox em versões anteriores. Para obter mais informações, consulte Versões de lançamento 4.x do Amazon EMR.

Se você usar uma AMI do Amazon Linux personalizada com base em uma AMI do Amazon Linux com data de criação 11/8/2018, o servidor Oozie falhará ao iniciar. Se você usar o Oozie, crie uma AMI personalizada com base em um ID de AMI do Amazon Linux com uma data de criação diferente. Use o AWS CLI comando da a seguir para retornar uma lista de imagens IDs de todas as imagens de HVM do Amazon Linux AMIs com a versão 2018.03, com a data de lançamento, para poder escolher uma AMI do Amazon Linux apropriada como base. MyRegion Substitua pelo identificador de sua região, como us-west-2.

```
aws ec2 --region MyRegion describe-images --owner amazon --query 'Images[?
Name!=`null`]|[?starts_with(Name, `amzn-ami-hvm-2018.03`) == `true`].
[CreationDate,ImageId,Name]' --output text | sort -rk1
```

A tabela a seguir lista a versão do Oozie incluída na versão mais recente da série 7.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Oozie.

Para obter a versão dos componentes instalados com o Oozie nessa versão, consulte Release 7.9.0 Component Versions.

Informações de versão do Oozie para o emr-7.9.0

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-7.9.0	Oozie 5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

A tabela a seguir lista a versão do Oozie incluída na versão mais recente da série 6.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Oozie.

Para obter a versão dos componentes instalados com o Oozie nessa versão, consulte Release 6.15.0 Component Versions.

Informações de versão do Oozie para o emr-6.15.0

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-6.15.0	Oozie 5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodema

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
		nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

A tabela a seguir lista a versão do Oozie incluída na versão mais recente da série 5.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Oozie.

Para obter a versão dos componentes instalados com o Oozie nesse lançamento, consulte Release 5.36.2 Component Versions.

Informações de versão do Oozie para o emr-5.36.2

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.36.2	Oozie 5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Tópicos

- Usar o Oozie com um banco de dados remoto no Amazon RDS
- Configurar a versão do Java para Oozie

- Histórico de versões do Oozie

Usar o Oozie com um banco de dados remoto no Amazon RDS

Por padrão, as informações de usuários do Oozie e os histórico de consultas são armazenados em um banco de dados MySQL local no nó principal. Como opção, é possível criar um ou mais clusters habilitados para o Oozie usando uma configuração armazenada no Amazon S3 e um banco de dados MySQL no Amazon Relational Database Service (Amazon RDS). Isso permite que você mantenha as informações dos usuários e o histórico de consultas criados pelo Oozie sem manter o cluster do Amazon EMR em execução. Recomendamos o uso da criptografia do lado do servidor do Amazon S3 para armazenar o arquivo de configuração.

Primeiro, crie o banco de dados remoto para o Oozie.

Para criar o banco de dados MySQL externo

1. Abra o console do Amazon RDS em https://console.aws.amazon.com/rds/.
2. Escolha Launch a DB Instance.
3. Selecione MySQL e, depois, escolha Select (Selecionar).
4. Deixe a seleção padrão de Multi-AZ Deployment and Provisioned IOPS Storage (Implantação Multi-AZ e armazenamento de IOPS provisionadas) e escolha Next (Próximo).
5. Deixe as especificações da instância em seus valores padrão, especifique configurações e selecione Next (Próximo).
6. Na página Configure Advanced Settings (Definir configurações avançadas), escolha um nome do grupo de segurança apropriado e o nome do banco de dados. O grupo de segurança utilizado deve permitir pelo menos o acesso TCP de entrada para a porta 3306 a partir do nó principal do seu cluster. Se você não tiver criado o cluster neste ponto, poderá permitir que todos os hosts se conectem à porta 3306 e ajustar o security group depois de ter executado o cluster. Escolha Launch DB Instance (Iniciar instância de banco de dados).
7. Em RDS Dashboard (Painel do RDS), selecione Instances (Instâncias) e selecione a instância que você acabou de criar. Quando seu banco de dados estiver disponível, anote o nome do banco de dados, o nome de usuário, a senha e o nome de host da instância do RDS. Você usará essas informações ao criar e configurar o cluster.

Para especificar um banco de dados MySQL externo para Oozie ao iniciar um cluster usando o AWS CLI

Para especificar um banco de dados MySQL externo para o Oozie ao executar um cluster com a AWS CLI, use as informações anotadas ao criar sua instância do RDS para configurar `oozie-site` com um objeto de configuração

> **Note**
>
> Você pode criar vários clusters que usam o mesmo banco de dados externo, mas cada cluster compartilhará informações de usuários e o histórico de usuários.

- Usando o AWS CLI, crie um cluster com o Oozie instalado, usando o banco de dados externo que você criou e referenciando um arquivo de configuração com uma classificação de configuração para o Oozie que especifica as propriedades do banco de dados. O exemplo a seguir cria um cluster com o Oozie instalado, fazendo referência a um arquivo de configuração no Amazon S3, `myConfig.json`, que especifica a configuração do banco de dados.

  > **Note**
  >
  > Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

  ```
  aws emr create-cluster --release-label emr-7.9.0 --applications Name=Oozie
   Name=Spark Name=Hive \
  --instance-type m5.xlarge --instance-count 3 \
  --configurations https://s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/
  myConfig.json --use-default-roles
  ```

  Os conteúdos de exemplo do arquivo `myConfig.json` são mostrados abaixo. Substitua *JDBC URL*, *username*, e *password* pelo URL do JDBC, nome de usuário e senha da sua instância do RDS.

> **Important**
>
> O URL do JDBC deve incluir o nome do banco de dados como um sufixo. Por exemplo, jdbc:mysql://.xxxxxxxxxx.us-east-1.rds.amazonaws.com:3306/ dbname. oozie-external-db

```
[{
  "Classification": "oozie-site",
    "Properties": {
        "oozie.service.JPAService.jdbc.driver": "org.mariadb.jdbc.Driver",
        "oozie.service.JPAService.jdbc.url": "JDBC URL",

        "oozie.service.JPAService.jdbc.username": "username",
        "oozie.service.JPAService.jdbc.password": "password"
    },
    "Configurations": []
}]
```

Configurar a versão do Java para Oozie

O Oozie executa vários processos da Java Virtual Machine (JVM). Esta página explica como configurar a versão do Java para cada processo.

- Servidor do Oozie: defina `JAVA_HOME` na classificação `oozie-env` para atualizar a versão do Java para o `EmbeddedOozieServer`.
- Oozie Launcher AM: Oozie Launcher AM é um trabalho de MR (MapReduce) de mapeador único que invoca as bibliotecas de clientes de aplicações apropriadas, como Hadoop e Hive. A menos que seja configurado de outra forma, as versões de runtime do Oozie Launcher AM são as mesmas dos runtimes do Java para Hadoop no cluster do EMR. Para configurar o runtime do Java para o Oozie Launcher AM, defina a seguinte propriedade no `workflow.xml` para o trabalho:

```
<property>
     <name>mapred.child.env</name>
     <value>JAVA_HOME=/path/to/JAVA_HOME</value>
 </property>
```

Essa propriedade garante que o Oozie Launcher AM para o trabalho do Oozie seja executado na versão do Java especificada por você, em vez de na versão do Java definida no Hadoop.

- Executável do cliente da aplicação: como o Oozie Launcher AM invoca o cliente da aplicação por padrão, o runtime do Java para o executável do cliente é o mesmo que o do Oozie Launcher AM.
- Aplicações iniciadas por um trabalho do Oozie: a menos que seja especificado de outra forma, as versões de runtime da aplicação real JVMs que são iniciadas por um trabalho do Oozie são as mesmas que as dos runtimes do Java para Hadoop no cluster do EMR. Dependendo do tipo de ação de fluxo de trabalho do Oozie usada para iniciar a aplicação em um trabalho do Oozie (ação do Spark ou do Hive), você pode atualizar o runtime padrão do Java para a aplicação real JVMs no trabalho do `workflow.xml` Oozie.

Histórico de versões do Oozie

A tabela a seguir informa a versão do Oozie incluída em cada versão do Amazon EMR e os componentes instalados com a aplicação. Para obter as versões dos componentes em cada versão, consulte a seção Versão de componentes da sua versão em Versões 7.x do Amazon EMR, Versões de lançamento 6.x do Amazon EMR ou Versões de lançamento 5.x do Amazon EMR.

Informações de versão do Oozie

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-7.9.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-7.8.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-7.7.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-7.6.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-7.5.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-7.4.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-7.3.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-7.2.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-5.36.2	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-7.1.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-7.0.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-6.15.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-6.14.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-6.13.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-6.12.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-6.11.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-6.11.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-6.10.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-6.10.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-6.9.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.9.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-6.8.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.8.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-6.7.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.36.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.36.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.6.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.35.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.5.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-6.4.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.3.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-6.3.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.2.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-6.2.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.1.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-6.1.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.0.1	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-6.0.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.34.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.33.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.33.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.32.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.32.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.31.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.31.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.30.2	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.30.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.30.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.29.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.28.1	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.28.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.27.1	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.27.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.26.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.25.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.24.1	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.24.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.23.1	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.23.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.22.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.21.2	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.21.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.21.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.20.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.20.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.19.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.19.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.18.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.18.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.17.2	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.17.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.17.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.16.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.16.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.15.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.15.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.14.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.14.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.14.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.13.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.13.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.12.3	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.12.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.12.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.12.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.11.4	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.11.3	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.11.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.11.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.11.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.10.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.10.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.9.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.9.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.8.3	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.8.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.8.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.8.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.7.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.7.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.6.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.6.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.5.4	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.5.3	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.5.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.5.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.5.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.4.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.4.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.3.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.3.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.3.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.2.3	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.2.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.2.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.2.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.1.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.1.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.0.3	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.0.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.0.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Oozie	Componentes instalados com o Oozie
emr-5.0.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Notas de lançamento do Oozie por versão

Tópicos

- Amazon EMR 7.6.0: alterações no Oozie
- Amazon EMR 7.5.0: alterações no Oozie
- Amazon EMR 7.4.0: alterações no Oozie
- Amazon EMR 7.3.0: alterações no Oozie
- Amazon EMR 7.1.0: alterações no Oozie
- Amazon EMR 7.0.0: alterações no Oozie

Amazon EMR 7.6.0: alterações no Oozie

Nenhuma alteração na versão 7.6.0.

Amazon EMR 7.5.0: alterações no Oozie

Tipo	Descrição
Upgrade	Remova o suporte para orquestrar trabalhos do Sqoop, pois o Sqoop não está disponível nas novas versões do EMR

Amazon EMR 7.4.0: alterações no Oozie

Mudanças no Oozie: SEM ALTERAÇÕES

Amazon EMR 7.3.0: alterações no Oozie

Tipo	Descrição
Upgrade	Guia de atualizações para a versão 4.0.
Alteração de comportamento padrão	Define duas propriedades em `oozie-site.xml` : • `oozie.service.SparkConfigurationService.spark.configurations=*=/etc/spark/conf/` • `oozie.service.SparkConfigurationService.spark.configurations.blacklist=spark.yarn.dist.files` Essas propriedades permitem que o Oozie escolha as configurações padrão do Spark no Amazon EMR ao instalar o Oozie com o Spark.

Amazon EMR 7.1.0: alterações no Oozie

Tipo	Descrição
Upgrade	Atualiza o Netty para 4.1.100.Final.
Upgrade	Atualiza o Jetty para 9.4.53.v20231009.

Amazon EMR 7.0.0: alterações no Oozie

Tipo	Descrição
Upgrade	O Oozie Client and Server agora usam o Java 17 por padrão. Consulte o Guia de versão do EMR 7.0.0 para obter mais detalhes.

Apache Phoenix

O Apache Phoenix é usado para o processamento de transações online (OLTP) e o analytics operacional, permitindo que você use consultas de SQL padrão e JDBC APIs para utilizar o armazenamento de backup do Apache. HBase Para obter mais informações, consulte Phoenix em 15 minutos ou menos. O Phoenix é incluído nas versões 4.7.0 e posteriores do Amazon EMR.

Se você fez a atualização de uma versão mais antiga do Amazon EMR para a versão 5.4.0 ou posterior do Amazon EMR e usa indexação secundária, atualize os índices locais conforme descrito na documentação do Apache Phoenix. O Amazon EMR remove as configurações necessárias da classificação do `hbase-site`, mas os índices precisam ser preenchidos novamente. O sistema oferece suporte a atualizações de índices online e offline. As atualizações online são o padrão, o que significa que os índices são preenchidos novamente durante a inicialização de clientes do Phoenix versão 4.8.0 ou posterior. Para especificar as atualizações offline, defina a configuração do `phoenix.client.localIndexUpgrade` como falsa na classificação do `phoenix-site` e, em seguida, execute o SSH no nó principal para executar o `psql [zookeeper] -1`.

A tabela a seguir lista a versão do Phoenix incluída na versão mais recente da série 7.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Phoenix.

Para obter a versão dos componentes instalados com o Phoenix nessa versão, consulte Release 7.9.0.

Informações de versão do Phoenix para o emr-7.9.0

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-7.9.0	Phoenix 5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
		timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-queryserver, zookeeper-client, zookeeper-server

A tabela a seguir lista a versão do Phoenix incluída na versão mais recente da série 6.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Phoenix.

Para obter a versão dos componentes instalados com o Phoenix nessa versão, consulte Release 6.15.0 Component Versions.

Informações de versão do Phoenix para o emr-6.15.0

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-6.15.0	Phoenix 5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
		server, zookeeper-client, zookeeper-server

A tabela a seguir lista a versão do Phoenix incluída na versão mais recente da série 5.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Phoenix.

Para obter a versão dos componentes instalados com o Phoenix nesse lançamento, consulte Release 5.36.2 Component Versions.

Informações de versão do Phoenix para o emr-5.36.2

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.36.2	Phoenix 4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Tópicos

- Criar um cluster com o Phoenix
- Clientes do Phoenix

- Histórico de versões do Phoenix

Criar um cluster com o Phoenix

Para instalar o Phoenix, escolha o aplicativo ao criar um cluster usando o console ou a AWS CLI. Os procedimentos e exemplos a seguir mostram como criar um cluster com o Phoenix HBase e. Para obter mais informações sobre como criar clusters usando o console, incluindo as Opções avançadas, consulte Plan and configure clusters no Guia de gerenciamento do Amazon EMR.

Para iniciar um cluster com o Phoenix instalado usando as Quick Options (Opções rápidas) para criar um cluster no console

1. Abra o console do Amazon EMR em https://console.aws.amazon.com /emr.
2. Escolha Create cluster (Criar cluster) para usar Quick Create.
3. Em Software Configuration (Configuração do software), escolha a versão mais recente apropriada para o seu aplicativo. O Phoenix aparece como uma opção somente quando a versão Amazon Release Version emr-4.7.0 ou posterior é selecionada.
4. Para Aplicativos, escolha a segunda opção HBase: HBase ***ver*** com Ganglia, Hadoop***ver***, Hive***ver***, Hue, ***ver*** Phoenix e. ***ver*** ***ver*** ZooKeeper ***ver***
5. Selecione outras opções conforme necessário e escolha Create cluster (Criar cluster).

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

O exemplo a seguir inicia um cluster com o Phoenix instalado usando as definições de configuração padrão.

Para iniciar um cluster com o Phoenix e HBase usar o AWS CLI

- Crie o cluster usando o seguinte comando:

```
aws emr create-cluster --name "Cluster with Phoenix" --release-label emr-7.9.0 \
```

```
--applications Name=Phoenix Name=HBase --ec2-attributes KeyName=myKey \
--instance-type m5.xlarge --instance-count 3 --use-default-roles
```

Personalizar as configurações do Phoenix

Ao criar um cluster, você configurar o Phoenix definindo os valores em `hbase-site.xml` usando a classificação da configuração `hbase-site`.

Para obter mais informações, consulte Configuração e ajustes na documentação do Phoenix.

O exemplo a seguir demonstra o uso de um arquivo JSON armazenado no Amazon S3 para especificar o valor `false` para a propriedade `phoenix.schema.dropMetaData`. Várias propriedades podem ser especificadas para uma única classificação. Para obter mais informações, consulte Configurar aplicações. O comando `create-cluster` faz referência ao arquivo JSON como o parâmetro `--configurations`.

O conteúdo do arquivo JSON salvo em to /amzn-s3-demo-bucket/myfolder/myconfig .json é mostrado a seguir.

```
[
    {
      "Classification": "hbase-site",
      "Properties": {
        "phoenix.schema.dropMetaData": "false"
      }
    }
  ]
```

O comando `create cluster` que faz referência ao arquivo JSON é mostrado no exemplo a seguir.

```
aws emr create-cluster --release-label emr-7.9.0 --applications Name=Phoenix \
Name=HBase --instance-type m5.xlarge --instance-count 2 \
--configurations https://s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myconfig.json
```

> **Note**
>
> A solicitação de reconfiguração para quaisquer classificações de configuração do Phoenix só é compatível nas versões 5.23.0 e posteriores do Amazon EMR, e não é compatível com a

versão 5.21.0 ou 5.22.0 do Amazon EMR. Para obter mais informações, consulte Fornecer uma configuração para um grupo de instâncias em um cluster em execução

Clientes do Phoenix

Você se conecta ao Phoenix usando um cliente do JDBC criado com dependências totais ou usando o "thin client", que usa o Phoenix Query Server e só pode ser executado em um nó principal do cluster (por exemplo, usando um cliente do SQL, uma etapa, uma linha de comando, um encaminhamento de porta de SSH, etc.). Ao usar o cliente JDBC “gordo”, ele ainda precisa ter acesso a todos os nós do cluster porque se conecta diretamente aos HBase serviços. O cliente "magro" do Phoenix precisa apenas do acesso ao Phoenix Query Server em uma porta padrão 8765. Há vários scripts no Phoenix que usam esses clientes.

Use uma etapa do Amazon EMR para fazer consultas usando o Phoenix

O procedimento a seguir restaura um instantâneo HBase e usa esses dados para executar uma consulta Phoenix. Você pode estender este exemplo ou criar um novo script que utiliza os clientes do Phoenix para atender às suas necessidades.

1. Crie um cluster com o Phoenix instalado usando o seguinte comando:

```
aws emr create-cluster --name "Cluster with Phoenix" --log-uri s3://amzn-s3-demo-bucket/myLogFolder --release-label emr-7.9.0 \
--applications Name=Phoenix Name=HBase --ec2-attributes KeyName=myKey \
--instance-type m5.xlarge --instance-count 3 --use-default-roles
```

2. Crie e, em seguida, carregue os seguintes arquivos para o Amazon S3:

   copySnapshot.sh

```
sudo su hbase -s /bin/sh -c 'hbase snapshot export \
 -D hbase.rootdir=s3://us-east-1.elasticmapreduce.samples/hbase-demo-customer-data/snapshot/ \
-snapshot customer_snapshot1 \
-copy-to hdfs://masterDNSName:8020/user/hbase \
-mappers 2 -chuser hbase -chmod 700'
```

   runQuery.sh

```
aws s3 cp s3://amzn-s3-demo-bucket/phoenixQuery.sql /home/hadoop/
/usr/lib/phoenix/bin/sqlline-thin.py http://localhost:8765 /home/hadoop/
phoenixQuery.sql
```

phoenixQuery.sql

> **Note**
>
> Você só precisa incluir `COLUMN_ENCODED_BYTES=0` no exemplo a seguir quando usar as versões 5.26.0 e posteriores do Amazon EMR.

```
CREATE VIEW "customer" (
pk VARCHAR PRIMARY KEY,
"address"."state" VARCHAR,
"address"."street" VARCHAR,
"address"."city" VARCHAR,
"address"."zip" VARCHAR,
"cc"."number" VARCHAR,
"cc"."expire" VARCHAR,
"cc"."type" VARCHAR,
"contact"."phone" VARCHAR)
COLUMN_ENCODED_BYTES=0;

CREATE INDEX my_index ON "customer" ("customer"."state") INCLUDE("PK",
 "customer"."city", "customer"."expire", "customer"."type");

SELECT "customer"."type" AS credit_card_type, count(*) AS num_customers FROM
 "customer" WHERE "customer"."state" = 'CA' GROUP BY "customer"."type";
```

Use o AWS CLI para enviar os arquivos para o bucket do S3:

```
aws s3 cp copySnapshot.sh s3://amzn-s3-demo-bucket/
aws s3 cp runQuery.sh s3://amzn-s3-demo-bucket/
aws s3 cp phoenixQuery.sql s3://amzn-s3-demo-bucket/
```

3. Crie uma tabela usando as seguintes etapas enviadas para o cluster que você criou na Etapa 1:

   createTable.json

```
[
  {
    "Name": "Create HBase Table",
    "Args": ["bash", "-c", "echo $'create \"customer\",\"address\",\"cc\",\"contact
\"' | hbase shell"],
    "Jar": "command-runner.jar",
    "ActionOnFailure": "CONTINUE",
    "Type": "CUSTOM_JAR"
  }
]
```

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps file://./createTable.json
```

4. Use o `script-runner.jar` para executar o script `copySnapshot.sh` que foi anteriormente carregado para o bucket do S3:

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps Type=CUSTOM_JAR,Name="HBase Copy Snapshot",ActionOnFailure=CONTINUE,\
Jar=s3://region.elasticmapreduce/libs/script-runner/script-
runner.jar,Args=["s3://amzn-s3-demo-bucket/copySnapshot.sh"]
```

   Isso executa um MapReduce trabalho para copiar seus dados de snapshot para o cluster HDFS.

5. Restaure o snapshot que você copiou para o cluster usando a seguinte etapa:

   restoreSnapshot.json

```
[
  {
    "Name": "restore",
    "Args": ["bash", "-c", "echo $'disable \"customer\"; restore_snapshot
 \"customer_snapshot1\"; enable \"customer\"' | hbase shell"],
    "Jar": "command-runner.jar",
    "ActionOnFailure": "CONTINUE",
    "Type": "CUSTOM_JAR"
  }
]
```

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
```

```
--steps file://./restoreSnapshot.json
```

6. Use o `script-runner.jar` para executar o script `runQuery.sh` que foi anteriormente carregado para o bucket do S3:

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps Type=CUSTOM_JAR,Name="Phoenix Run Query",ActionOnFailure=CONTINUE,\
Jar=s3://region.elasticmapreduce/libs/script-runner/script-
runner.jar,Args=["s3://amzn-s3-demo-bucket/runQuery.sh"]
```

   A consulta é executada e obtém os resultados do `stdout` da etapa. Pode levar alguns minutos para a etapa ser concluída.

7. Examine os resultados do `stdout` da etapa no log do URI que você usou quando criou o cluster na Etapa 1. Os resultados devem ser algo semelhante a:

```
+-------------------------------------------+-----------------------------------+
|             CREDIT_CARD_TYPE              |              NUM_CUSTOMERS        |
+-------------------------------------------+-----------------------------------+
| american_express                          | 5728                              |
| dankort                                   | 5782                              |
| diners_club                               | 5795                              |
| discover                                  | 5715                              |
| forbrugsforeningen                        | 5691                              |
| jcb                                       | 5762                              |
| laser                                     | 5769                              |
| maestro                                   | 5816                              |
| mastercard                                | 5697                              |
| solo                                      | 5586                              |
| switch                                    | 5781                              |
| visa                                      | 5659                              |
+-------------------------------------------+-----------------------------------+
```

Histórico de versões do Phoenix

A tabela a seguir informa a versão do Phoenix incluída em cada versão do Amazon EMR e os componentes instalados com a aplicação. Para obter as versões dos componentes em cada versão, consulte a seção Versão de componentes da sua versão em Versões 7.x do Amazon EMR, Versões de lançamento 6.x do Amazon EMR ou Versões de lançamento 5.x do Amazon EMR.

Informações sobre a versão do Phoenix

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-7.9.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-queryserver, zookeeper-client, zookeeper-server
emr-7.8.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-operator-tools, phoenix-library, phoenix-

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
		connectors, phoenix-query-server, zookeeper-client, zookeeper-server
emr-7.7.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-7.6.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-7.5.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-7.4.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-7.3.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-7.2.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.36.2	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-7.1.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-7.0.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-6.15.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-6.14.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-6.13.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-6.12.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-6.11.1	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-6.11.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-queryserver, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-6.10.1	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-6.10.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-6.9.1	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-6.9.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-6.8.1	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-6.8.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-6.7.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.36.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.36.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-6.6.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.35.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-6.5.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-6.4.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-6.3.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-6.3.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-6.2.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-6.2.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-6.1.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-6.1.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-6.0.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-6.0.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.34.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.33.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.33.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.32.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.32.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.31.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.31.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.30.2	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.30.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.30.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.29.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.28.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.28.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.27.1	4.14.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.27.0	4.14.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.26.0	4.14.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.25.0	4.14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.24.1	4.14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.24.0	4.14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.23.1	4.14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.23.0	4.14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.22.0	4.14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.21.2	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.21.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.21.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.20.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.20.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.19.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.19.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.18.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.18.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.17.2	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.17.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.17.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.16.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.16.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.15.1	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.15.0	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.14.2	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.14.1	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.14.0	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.13.1	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.13.0	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.12.3	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.12.2	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.12.1	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.12.0	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.11.4	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.11.3	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.11.2	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.11.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.11.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.10.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.10.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.9.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.9.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.8.3	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.8.2	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.8.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.8.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.7.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.7.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.6.1	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.6.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.5.4	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.5.3	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.5.2	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.5.1	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.5.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.4.1	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.4.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.3.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.3.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.3.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.2.3	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.2.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.2.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.2.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.1.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.1.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.0.3	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.0.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.0.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-5.0.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.9.6	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-4.9.5	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.9.4	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-4.9.3	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.9.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-4.9.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.8.5	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-4.8.4	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.8.3	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-4.8.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.8.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-4.8.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.7.4	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-4.7.3	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.7.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	Versão do Phoenix	Componentes instalados com o Phoenix
emr-4.7.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.7.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Apache Pig

O Apache Pig é uma biblioteca Apache de código aberto que é executada sobre o Hadoop, fornecendo uma linguagem de scripts que você pode usar para transformar grandes conjuntos de dados sem precisar gravar códigos complexos em uma linguagem de computação de nível inferior, como Java. A biblioteca usa comandos semelhantes a SQL escritos em uma linguagem chamada Pig Latin e converte esses comandos em tarefas Tez com base em gráficos acíclicos direcionados () ou programas. DAGs MapReduce O Pig funciona com dados estruturados e não estruturados em vários formatos. Para obter mais informações sobre o Pig, consulte http://pig.apache.org/.

Você pode executar comandos do Pig interativamente ou no modo em lote. Para usar o Pig interativamente, crie uma conexão SSH com o nó principal e envie os comandos usando o shell do Grunt. Para usar o Pig no modo em lote, escreva os scripts do Pig, carregue-os no Amazon S3 e envie-os como etapas do cluster. Para obter mais informações sobre como enviar trabalhos para um cluster, consulte Enviar trabalhos para um cluster no Guia de gerenciamento do Amazon EMR.

Ao usar o Pig para gravar a saída em uma HCatalog tabela no Amazon S3, desabilite a gravação direta do Amazon EMR `mapred.output.direct.NativeS3FileSystem` definindo as propriedades e como. `mapred.output.direct.EmrFileSystem` `false` Para obter mais informações, consulte Usando HCatalog. Dentro de um script do Pig, você pode usar os comandos `SET mapred.output.direct.NativeS3FileSystem false` e `SET mapred.output.direct.EmrFileSystem false`.

A tabela a seguir lista a versão do Pig incluída na versão mais recente da série 7.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Pig.

Para obter a versão dos componentes instalados com o Pig nesse lançamento, consulte Release 7.9.0 Component Versions.

Informações de versão do Pig para o emr-7.9.0

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-7.9.0	Pig 0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
		hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

A tabela a seguir lista a versão do Pig incluída na versão mais recente da série 6.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Pig.

Para obter a versão dos componentes instalados com o Pig nessa versão, consulte Release 6.15.0 Component Versions.

Informações de versão do Pig para o emr-6.15.0

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-6.15.0	Pig 0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

A tabela a seguir lista a versão do Pig incluída na versão mais recente da série 5.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Pig.

Para obter a versão dos componentes instalados com o Pig nesse lançamento, consulte Release 5.36.2 Component Versions.

Informações de versão do Pig para o emr-5.36.2

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.36.2	Pig 0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Tópicos

- Enviar um trabalho do Pig
- Chamar funções definidas pelo usuário no Pig
- Histórico de versões do Pig

Enviar um trabalho do Pig

Esta seção demonstra como enviar trabalhos do Pig para um cluster do Amazon EMR. Os exemplos a seguir geram um relatório contendo o total de bytes transferidos, uma lista dos 50 principais endereços IP, uma lista dos 50 principais indicadores externos e os 50 principais termos de pesquisa usando o Bing e o Google. O script do Pig está localizado no bucket `s3://elasticmapreduce/samples/pig-apache/do-reports2.pig` do Amazon S3. Os dados de entrada estão localizados no bucket `s3://elasticmapreduce/samples/pig-apache/input` do Amazon S3. Os dados de saída são salvos em um bucket do Amazon S3.

Enviar trabalhos do Pig usando o console do Amazon EMR

Este exemplo descreve como usar o console do Amazon EMR para adicionar uma etapa do Pig a um cluster.

Para enviar uma etapa do Pig

1. Abra o console do Amazon EMR em https://console.aws.amazon.com /emr.
2. Escolha Criar cluster para criar um cluster com o Pig instalado. Para conhecer as etapas sobre como criar um cluster, consulte Planejar e configurar um cluster do Amazon EMR.
3. Abra um terminal e use SSH no nó principal do cluster seguindo as etapas descritas em Conectar-se ao nó principal usando SSH. Depois de fazer isso, execute as etapas a seguir.

   ```
   sudo mkdir -p /home/hadoop/lib/pig/
   sudo aws s3 cp s3://elasticmapreduce/libs/pig/0.3/piggybank-0.3-amzn.jar /home/
   hadoop/lib/pig/piggybank.jar
   ```

4. No console, clique em Lista de clusters e selecione o nome do cluster que você criou.
5. Role até a seção Steps (Etapas) e expanda-a. Em seguida, escolha Add step (Adicionar etapa).
6. Na caixa de diálogo Add Step (Adicionar etapa):
   - Para Step type (Tipo de etapa), escolha Pig program (Programa Pig).
   - Para Name (Nome), aceite o nome padrão (programa Pig) ou digite um novo nome.
   - Para Script S3 location (Local do script do S3), digite o local do script do Pig. Por exemplo: **`s3://elasticmapreduce/samples/pig-apache/do-reports2.pig`**.
   - Para Input S3 location (Local de entrada do S3), digite o local dos dados de entrada. Por exemplo: **`s3://elasticmapreduce/samples/pig-apache/input`**.
   - Em Local de saída do S3, digite ou navegue até o nome do bucket de saída do Amazon S3.
   - Para Arguments (Argumentos), deixe o campo em branco.
   - Para Action on failure (Ação na falha), aceite a opção padrão Continue (Continuar).
7. Escolha Adicionar. A etapa é exibida no console com o status Pendente.
8. O status da etapa muda de Pending (Pendente) para Running (Em execução) e depois para Completed (Concluído) conforme ela é executada. Para atualizar o status, escolha o ícone Refresh (Atualizar) acima da coluna Actions (Ações). Quando sua etapa estiver concluída, verifique o bucket do Amazon S3 para confirmar se os arquivos de saída da etapa do Pig estão lá.

Enviar trabalhos do Pig usando a AWS CLI

Para enviar uma etapa do Pig usando o AWS CLI

Ao iniciar um cluster usando o AWS CLI, use o `--applications` parâmetro para instalar o Pig. Para enviar uma etapa do Pig, use o parâmetro `--steps`.

1. Para iniciar um cluster com o Pig instalado, digite o comando a seguir, substituindo *myKey* e *amzn-s3-demo-bucket/* pelo nome do par de EC2 chaves e do bucket do Amazon S3.

```
aws emr create-cluster \
--name "Test cluster" \
--log-uri s3://amzn-s3-demo-bucket/ \
--release-label emr-5.36.2 \
--applications Name=Pig \
--use-default-roles \
--ec2-attributes KeyName=myKey \
--instance-type m5.xlarge \
--instance-count 3
```

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

Quando você especifica a contagem de instâncias sem usar o parâmetro `--instance-groups`, um único nó principal é executado, e as instâncias restantes são executadas como nós core. Todos os nós usam o tipo de instância especificado no comando.

> **Note**
>
> Se você não tiver criado o perfil de serviço padrão do EMR e o perfil de EC2 instância, digite `aws emr create-default-roles` para criá-los antes de digitar o subcomando. `create-cluster`

2. Para enviar uma etapa do Pig, digite o comando a seguir, substituindo *myClusterId* e *amzn-s3-demo-bucket* pelo ID do cluster e pelo nome do bucket do Amazon S3.

```
aws emr add-steps \
--cluster-id myClusterId \
--steps Type=PIG,Name="Pig Program",ActionOnFailure=CONTINUE,Args=[-f,s3://
elasticmapreduce/samples/pig-apache/do-reports2.pig,-p,INPUT=s3://elasticmapreduce/
samples/pig-apache/input,-p,OUTPUT=s3://amzn-s3-demo-bucket/pig-apache/output]
```

Esse comando retornará um ID de etapa, que você pode usar para verificar o `State` da sua etapa.

3. Consulte o status da sua etapa com o comando `describe-step`.

```
aws emr describe-step --cluster-id myClusterId --step-id s-1XXXXXXXXXXA
```

O `State` da etapa é alterado de `PENDING` para `RUNNING` e para `COMPLETED`, conforme a etapa é executada. Quando sua etapa estiver concluída, verifique o bucket do Amazon S3 para confirmar se os arquivos de saída da etapa do Pig estão lá.

Para obter mais informações sobre o uso de comandos do Amazon EMR na AWS CLI, consulte AWS CLI Command Reference.

Chamar funções definidas pelo usuário no Pig

O Pig fornece a capacidade de chamar funções definidas pelo usuário (UDFs) de dentro dos scripts do Pig. Você pode fazer isso para implementar o processamento personalizado em seus scripts do Pig. As linguagens atualmente suportadas são Java, Python/Jython e JavaScript (embora JavaScript o suporte ainda seja experimental).

As seções a seguir descrevem como registrar suas funções no Pig para que você possa chamá-las a partir do shell ou de scripts do Pig. Para obter mais informações sobre o uso UDFs com o Pig, consulte a Documentação do Pig da sua versão do Pig.

Chamada de arquivos JAR a partir do Pig

Você pode usar arquivos JAR personalizados com o Pig usando o comando `REGISTER` no seu script do Pig. O arquivo JAR é local ou um sistema de arquivos remoto, como o Amazon S3. Quando o script do Pig é executado, o Amazon EMR baixa o arquivo JAR automaticamente para o nó principal e, em seguida, carrega o arquivo JAR no cache distribuído do Hadoop. Dessa forma, o arquivo JAR é automaticamente usado, quando necessário, por todas as instâncias no cluster.

Para usar arquivos JAR com o Pig

1. Carregue o arquivo JAR personalizado no Amazon S3.
2. Use o comando `REGISTER` no script do Pig para especificar o bucket no Amazon S3 do arquivo JAR personalizado.

```
REGISTER s3://amzn-s3-demo-bucket/path/mycustomjar.jar;
```

Chamar scripts em Python/Jython no Pig

Você pode registrar scripts em Python no Pig e, em seguida, chamar funções nesses scripts a partir do shell do Pig ou de um script do Pig. Para fazer isso, especifique o local do script com a palavra-chave `register`.

Como o Pig em escrito em Java, ele usa o mecanismo de script Jython para analisar scripts do Python. Para obter mais informações sobre Jython, acesse http://www.jython.org/.

Para chamar scripts em Python/Jython a partir do Pig

1. Escreva um script em Python e carregue-o para um local no Amazon S3. Esse deve ser um bucket de propriedade da mesma conta que estiver criando o cluster do Pig, ou um bucket que tenha permissões definidas de tal forma que a conta que criou o cluster pode acessá-lo. Neste exemplo, o upload do script é feito em s3://amzn-s3-demo -bucket/pig/python.
2. Inicie um cluster Pig. Se você estiver acessando o Pig a partir do shell Grunt, execute um cluster interativo. Se você estiver executando comandos do Pig em um script, inicie um cluster Pig com script. Este exemplo inicia um cluster interativo. Para obter mais informações sobre como criar um cluster Pig, consulte Enviar um trabalho do Pig.
3. Para um cluster interativo, use o SSH para conectar-se ao nó principal e executar o shell Grunt. Para obter mais informações, consulte SSH para o nó principal.
4. Execute o shell Grunt do Pig, digitando `pig` na linha de comando:

```
pig
```

5. Registre a biblioteca Jython e o script do Python no Pig usando a palavra-chave `register` no prompt de comando do Grunt, como mostrado no comando a seguir, no qual você deve especificar o local do seu script no Amazon S3:

```
grunt> register 'lib/jython.jar';
grunt> register 's3://amzn-s3-demo-bucket/pig/python/myscript.py' using jython as
 myfunctions;
```

6. Carregue os dados de entrada. O exemplo a seguir carrega a entrada de um local do Amazon S3:

```
grunt> input = load 's3://amzn-s3-demo-bucket/input/data.txt' using TextLoader as
 (line:chararray);
```

7. Agora, você pode chamar funções no seu script dentro Pig, fazendo referência a eles usando `myfunctions`:

```
grunt> output=foreach input generate myfunctions.myfunction($1);
```

Histórico de versões do Pig

A tabela a seguir informa a versão do Pig incluída em cada versão do Amazon EMR e os componentes instalados com a aplicação. Para obter as versões dos componentes em cada versão, consulte a seção Versão de componentes da sua versão em Versões 7.x do Amazon EMR, Versões de lançamento 6.x do Amazon EMR ou Versões de lançamento 5.x do Amazon EMR.

Informações de versão do Pig

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-7.9.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
		timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-7.8.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-7.7.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-7.6.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-7.5.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-7.4.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-7.3.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-7.2.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-5.36.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-7.1.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-7.0.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-6.15.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-6.14.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-6.13.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-6.12.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-6.11.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-6.11.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-6.10.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-6.10.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-6.9.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.9.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-6.8.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.8.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-6.7.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.36.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.36.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.6.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.35.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.5.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-6.4.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.3.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-6.3.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.2.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-6.2.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.1.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-6.1.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.34.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.33.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.33.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.32.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.32.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.31.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.31.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.30.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.30.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.30.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.29.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.28.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.28.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.27.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.27.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.26.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.25.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.24.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.24.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.23.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.23.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.22.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.21.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.21.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.21.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.20.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.20.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.19.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.19.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.18.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.18.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.17.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.17.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.17.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.16.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.16.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.15.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.15.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.14.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.14.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.14.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.13.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.13.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.12.3	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.12.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.12.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.12.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.11.4	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.11.3	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.11.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.11.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.11.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.10.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.10.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.9.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.9.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.8.3	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.8.2	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.8.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.8.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.7.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.7.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.6.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.6.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.5.4	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.5.3	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.5.2	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.5.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.5.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.4.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.4.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.3.2	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.3.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.3.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.2.3	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.2.2	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.2.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.2.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.1.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.1.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.0.3	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.0.2	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.0.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-5.0.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-4.9.6	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.9.5	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-4.9.4	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.9.3	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.9.2	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-4.9.1	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.8.5	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.8.4	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-4.8.3	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.8.2	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.8.1	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-4.8.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.7.4	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.7.3	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-4.7.2	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.7.1	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.7.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-4.6.1	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.6.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.5.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-4.4.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.3.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.2.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Rótulo de versão do Amazon EMR	Versão do Pig	Componentes instalados com o Pig
emr-4.1.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.0.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Presto

> **Important**
>
> O Presto, a versão anterior do Trino, ainda está disponível para uso com o Amazon EMR. No entanto, é altamente recomendável usar o Trino no futuro com o Amazon EMR. Observe também que o Trino e o Presto não podem ser executados simultaneamente no mesmo cluster. Para obter mais informações, consulte Trino.

O Presto é um mecanismo de consulta SQL rápido, projetado para consultas analíticas interativas em grandes conjuntos de dados de várias fontes. Para obter mais informações, consulte o Site do Presto. O Presto está incluído nas versões 5.0.0 e posteriores do Amazon EMR. As versões anteriores incluem o Presto como um aplicação em sandbox. Para obter mais informações, consulte Versões de lançamento 4.x do Amazon EMR. As versões 6.1.0 e posteriores do Amazon EMR são compatíveis com o Trino, além do Presto. Para obter mais informações, consulte Instalação do PrestoDB.

A tabela a seguir lista a versão do Presto incluída na versão mais recente da série 7.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Presto.

Para obter a versão dos componentes instalados com o Presto nessa versão, consulte Release 7.9.0 Component Versions.

Informações de versão do Presto para o emr-7.9.0

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
amzn-7.9.0	Presto 0.287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
		client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

A tabela a seguir lista a versão do Presto incluída na versão mais recente da série 6.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Presto.

Para obter a versão dos componentes instalados com o Presto nessa versão, consulte Release 6.15.0 Component Versions.

Informações de versão do Presto para o emr-6.15.0

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-6.15.0	Presto 0.283	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

A tabela a seguir lista a versão do Presto incluída na versão mais recente da série 5.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Presto.

Para obter a versão dos componentes instalados com o Presto nesse lançamento, consulte Release 5.36.2 Component Versions.

Informações de versão do Presto para o emr-5.36.2

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.36.2	Presto 0.267	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Tópicos

- Usar o Presto com o Catálogo de Dados do AWS Glue
- Usar S3 Select Pushdown com o Presto para melhorar a performance
- Adicionar conectores de banco de dados
- Usar SSL/TLS e configurar LDAPS com o Presto no Amazon EMR
- Ativar o modo estrito do Presto
- Lidar com a perda de instâncias spot no Presto
- Usar a escalabilidade automática do Presto com desativação tranquila
- Considerações com o Presto no Amazon EMR
- Histórico de versões do Presto

Usar o Presto com o Catálogo de Dados do AWS Glue

Ao usar as versões 5.10.0 e posteriores do Amazon EMR, você pode especificar o Catálogo de Dados do AWS Glue como o metastore padrão do Hive para o Presto. Recomendamos essa

configuração quando você precisa de um metastore persistente ou de um metastore compartilhado por diferentes clusters, serviços, aplicações ou Contas da AWS.

AWS O Glue é um serviço de extração, transformação e carregamento (ETL) totalmente gerenciado que torna fácil e econômico categorizar os dados, limpá-los, aprimorá-los e movê-los de modo confiável entre vários armazenamentos de dados. O Catálogo de Dados do AWS Glue fornece um repositório de metadados unificado em uma grande variedade de fontes e formatos de dados, integrando-se não apenas ao Amazon EMR, mas ao Amazon RDS, ao Amazon Redshift, ao Redshift Spectrum, ao Athena e a qualquer aplicação compatível com o metastore do Apache Hive. Os crawlers do AWS Glue podem inferir automaticamente o esquema com base nos dados de origem no Amazon S3 e armazenar os metadados associados no Catálogo de Dados. Para obter mais informações sobre o Catálogo de Dados, consulte Preencher o Catálogo de Dados do AWS Glue no Guia do desenvolvedor do AWS Glue.

São feitas cobranças separadas pelo uso do AWS Glue. Existe uma taxa mensal para armazenar e acessar os metadados no Catálogo de Dados, uma taxa horária cobrada por minuto para trabalhos de extração, transformação e carregamento (ETL) do AWS Glue e runtime do crawler e uma taxa horária cobrada por minuto para cada endpoint de desenvolvimento provisionado. O Catálogo de Dados permite o armazenamento de até um milhão de objetos gratuitamente. Se você armazenar mais de um milhão de objetos, será cobrado uma taxa de USD$1 por cada 100.000 objetos em mais de um milhão. Um objeto no Catálogo de Dados é uma tabela, uma partição ou um banco de dados. Para obter mais informações, consulte Definição de preço do Glue.

> **Important**
>
> Se você tiver criado tabelas usando o Amazon Athena ou o Amazon Redshift Spectrum antes de 14 de agosto de 2017, os bancos de dados e as tabelas serão armazenados em um catálogo gerenciado pelo Athena, que é separado do Catálogo de Dados do Glue. AWS Para integrar o Amazon EMR a essas tabelas, faça a atualização para o Catálogo de Dados do AWS Glue. Para obter mais informações, consulte Atualização para o catálogo de dados AWS Glue no Guia do usuário do Amazon Athena.

Especificar o Catálogo de Dados do AWS Glue como metastore

Você pode especificar o Catálogo de Dados do AWS Glue como metastore usando o AWS Management Console AWS CLI, a ou a API do Amazon EMR. Ao usar a CLI ou a API, você usa a classificação de configuração do Presto para especificar o Catálogo de Dados. Além disso, com

as versões 5.16.0 e posteriores do Amazon EMR, você pode usar a classificação de configuração para especificar um Catálogo de Dados em outra. Conta da AWS Ao usar o console, você pode especificar o Catálogo de Dados usando Opções avançadas ou Opções rápidas.

Console

Especificar o Catálogo de Dados do AWS Glue como metastore do Hive com o novo console

1. Faça login no e abra AWS Management Console o console do Amazon EMR em https://console.aws.amazon.com /emr.
2. Em EMR, EC2 no painel de navegação esquerdo, escolha Clusters e, em seguida, escolha Criar cluster.
3. Em Pacote de aplicações, escolha Presto.
4. Em Configurações do Catálogo de Dados do AWS Glue, marque a caixa de seleção Usar para metadados de tabelas do Presto.
5. Escolha qualquer outra opção que se aplique ao cluster.
6. Para iniciar o cluster, escolha Criar cluster.

CLI

Especificar o Catálogo de Dados do AWS Glue como o metastore padrão do Hive usando a AWS CLI

Para obter exemplos de como especificar as classificações de configuração a seguir ao criar um cluster, consulte Configurar aplicações.

Versões 5.16.0 e posteriores do Amazon EMR

- Defina a propriedade `hive.metastore` como `glue`, conforme mostrado no exemplo JSON a seguir.

```
[
  {
    "Classification": "presto-connector-hive",
    "Properties": {
      "hive.metastore": "glue"
    }
  }
]
```

Para especificar um Catálogo de Dados em outra conta da Conta da AWS, adicione a `hive.metastore.glue.catalogid` propriedade conforme mostrado no exemplo de JSON a seguir. Substitua *`acct-id`* pela conta da AWS do Catálogo de Dados. O uso de um Catálogo de Dados em outra Conta da AWS não está disponível nas versões 5.15.0 e anteriores do Amazon EMR.

```
[
  {
    "Classification": "presto-connector-hive",
    "Properties": {
      "hive.metastore": "glue",
      "hive.metastore.glue.catalogid": "acct-id"
    }
  }
]
```

Amazon EMR 5.10.0 a 5.15.0

Defina a propriedade `hive.metastore.glue.datacatalog.enabled` como `true`, conforme mostrado no exemplo JSON a seguir:

```
[
  {
    "Classification": "presto-connector-hive",
    "Properties": {
      "hive.metastore.glue.datacatalog.enabled": "true"
    }
  }
]
```

Versões 6.1.0 e posteriores do Amazon EMR usando PrestoSQL (Trino)

Desde a versão 6.1.0 do EMR, o PrestoSQL também é compatível com o Glue como o metastore padrão do Hive. Use a classificação de configuração `prestosql-connector-hive` e defina a propriedade `hive.metastore` como `glue`, conforme mostrado no exemplo de JSON a seguir.

As versões 6.4.0 e posteriores do Amazon EMR usam o novo nome Trino em vez de PrestoSQL. Se você usa o Trino, substitua *prestosql-connector-hive* na seguinte classificação de configuração por `trino-connector-hive`.

```
[
  {
    "Classification": "prestosql-connector-hive",
    "Properties": {
      "hive.metastore": "glue"
    }
  }
]
```

Para alternar entre metastores em um cluster de execução prolongada, você pode definir esses valores manualmente, conforme apropriado para a sua versão, conectando-se ao nó principal, editando os valores de propriedade diretamente no arquivo `/etc/presto/conf/catalog/hive.properties` e reiniciando o servidor Presto (`sudo restart presto-server`). Se você usar esse método com o as versões 5.15.0 e anteriores do Amazon EMR, certifique-se de que `hive.table-statistics-enabled` esteja definido como `false`. Essa configuração não é necessária ao usar versões 5.16.0 e posteriores. No entanto, as estatísticas de tabela e de partição não são compatíveis.

permissões do IAM

O perfil de EC2 instância de um cluster deve ter permissões do IAM para ações do AWS Glue. Além disso, se você habilitar a criptografia para objetos do Catálogo de Dados do AWS Glue, o perfil também deverá ter permissão para criptografar, descriptografar e gerar a AWS KMS key usada para criptografia.

Permissões para todas as ações do AWS Glue

Se você usa o perfil de EC2 instância padrão para o Amazon EMR, nenhuma ação será necessária. A política `AmazonElasticMapReduceforEC2Role` gerenciada anexada ao `EMR_EC2_DefaultRole` permite todas as ações necessárias do AWS Glue. No entanto, se você especificar um perfil e permissões de EC2 instância personalizados, deverá configurar as ações apropriadas do AWS Glue. Use a política gerenciada `AmazonElasticMapReduceforEC2Role`

como ponto de partida. Para obter mais informações, consulte Função de serviço para EC2 instâncias de cluster (perfil de EC2 instância) no Guia de gerenciamento do Amazon EMR.

Permissões para criptografar e descriptografar o Catálogo de Dados do AWS Glue

O perfil de instância precisa de permissão para criptografar e descriptografar dados usando a chave. Você não precisa configurar essas permissões se as duas instruções a seguir se aplicarem:

- Você habilita a criptografia para objetos do Catálogo de Dados do AWS Glue usando chaves gerenciadas para o AWS Glue.
- Você usa um cluster que está na mesma que o Catálogo Conta da AWS de Dados do AWS Glue.

Caso contrário, você deverá adicionar a seguinte instrução à política de permissões anexada ao perfil de EC2 instância.

```
[
    {
        "Version": "2012-10-17",
        "Statement": [
            {
                "Effect": "Allow",
                "Action": [
                    "kms:Decrypt",
                    "kms:Encrypt",
                    "kms:GenerateDataKey"
                ],
                "Resource": "arn:aws:kms:region:acct-
id:key/12345678-1234-1234-1234-123456789012"
            }
        ]
    }
]
```

Para obter mais informações sobre criptografia do Catálogo de Dados do AWS Glue, consulte Criptografar seu catálogo de dados no Guia do desenvolvedor do AWS Glue.

Permissões baseadas em recursos

Se você usa o AWS Glue com o Hive, o Spark ou o Presto no Amazon EMR AWS , o Glue oferece suporte a políticas baseadas em recursos para controlar o acesso aos recursos do Catálogo de Dados. Esses recursos incluem bancos de dados, tabelas, conexões e funções definidas pelo

usuário. Para obter mais informações, consulte Políticas baseadas em recursos no AWS Glue no Guia do desenvolvedor do AWS Glue.

Ao usar políticas baseadas em recursos para limitar o acesso ao AWS Glue de dentro do Amazon EMR, o principal que você especifica na política de permissões deve ser o ARN do perfil associado ao perfil de EC2 instância que é especificado quando um cluster é criado. Por exemplo, para uma política baseada em recursos anexada a um catálogo, você pode especificar o ARN do perfil para o perfil de serviço padrão de EC2 instâncias do cluster, *EMR_EC2_DefaultRole* como a`Principal`, usando o formato mostrado no seguinte exemplo:

```
arn:aws:iam::acct-id:role/EMR_EC2_DefaultRole
```

*acct-id*Pode ser diferente do ID da conta do AWS Glue. Isso possibilita o acesso de clusters do EMR em outras contas. Você pode especificar várias entidades principais, cada uma de uma conta diferente.

Considerações ao usar o Catálogo de Dados do AWS Glue

Considere os seguintes itens ao usar o Catálogo de Dados do AWS Glue como um metastore com o Presto:

- Não há suporte para renomear tabelas de dentro do AWS Glue.
- Quando você cria uma tabela do Hive sem especificar um `LOCATION`, os dados da tabela são armazenados no local especificado pela propriedade `hive.metastore.warehouse.dir`. Por padrão, esse é um local no HDFS. Se outro cluster precisar acessar a tabela, ele apresentará falha, a menos que tenha permissões adequadas para o cluster que criou a tabela. Além disso, como o armazenamento do HDFS é transitório, se o cluster for encerrado, os dados da tabela serão perdidos e a tabela precisará ser recriada. Recomendamos que você especifique a `LOCATION` no Amazon S3 ao criar uma tabela do Hive usando o Glue. AWS Como alternativa, você pode usar a classificação de configuração `hive-site` para especificar um local no Amazon S3 para `hive.metastore.warehouse.dir`, que se aplica a todas as tabelas do Hive. Se uma tabela for criada em um local do HDFS e o cluster que a criou ainda estiver em execução, você poderá atualizar o local da tabela para o Amazon S3 de dentro AWS do Glue. Para obter mais informações, consulte Como trabalhar com tabelas no console do AWS Glue no Guia do desenvolvedor do AWS Glue.
- Valores de partição contendo aspas e apóstrofos não são compatíveis, por exemplo, `PARTITION (owner="Doe's").`

- As estatísticas de coluna são têm suporte para o emr-5.31.0 e versões posteriores.
- O uso da autorização do Hive não é compatível. Como alternativa, considere usar as Políticas baseadas em recursos do AWS Glue. Para obter mais informações, consulte Usar políticas baseadas em recursos para acesso do Amazon EMR ao Catálogo de Dados do AWS Glue.

Usar S3 Select Pushdown com o Presto para melhorar a performance

> **Important**
>
> O Amazon S3 Select não está mais disponível para novos clientes. Os clientes existentes do Amazon S3 Select podem continuar usando o recurso normalmente. Saiba mais

Com as versões 5.18.0 e posteriores do Amazon EMR, você pode usar S3 Select Pushdown com Presto no Amazon EMR. Esse atributo permite que o Presto "empurre" o trabalho de computação de operações de projeção (por exemplo, `SELECT`) e operações de predicado (por exemplo, `WHERE`) para o Amazon S3. Isso permite que as consultas recuperem apenas os dados necessários do Amazon S3, o que pode melhorar a performance e reduzir a quantidade de dados transferidos entre o Amazon EMR e o Amazon S3 em algumas aplicações.

O S3 Select Pushdown é adequado para minha aplicação?

Recomendamos que você avalie seus aplicativos com e sem o S3 Selecione Pushdown para ver se o uso pode ser adequado para o aplicativo.

Use as seguintes diretrizes para determinar se o seu aplicativo é adequado para o uso do S3 Select:

- Sua consulta filtra mais de metade do conjunto de dados original.
- Os predicados do filtro de consulta usam colunas que têm um tipo de dados compatível com o Presto e o S3 Select. Os tipos de dados real e duplo e o carimbo de data e hora não têm suporte no S3 Select Pushdown. É recomendável usar o tipo de dados decimal para dados numéricos. Para obter mais informações sobre tipos de dados compatíveis com o S3 Select, consulte Tipos de dados no Guia do usuário do Amazon Simple Storage Service.
- Sua conexão de rede entre o Amazon S3 e o cluster do Amazon EMR tem boa velocidade de transferência e largura de banda disponível. O Amazon S3 não compacta respostas HTTP. Portanto, é provável que o tamanho da resposta aumente para arquivos de entrada compactados.

Considerações e limitações

- Somente objetos armazenados em formato CSV têm suporte. Os objetos podem ser descompactados ou compactados com gzip ou bzip2.
- A propriedade `AllowQuotedRecordDelimiters` não é compatível. Se essa propriedade for especificada, a consulta falhará.
- A criptografia do lado do servidor do Amazon S3 com chaves de criptografia fornecidas pelo cliente (SSE-C) e a criptografia do lado do cliente não são compatíveis.
- O S3 Select Pushdown não é um substituto para o uso de formatos de arquivo colunar ou compactado, como ORC ou Parquet.
- O Amazon S3 Select não oferece suporte a consultas entre regiões para o Presto e o Trino.

Habilitar o S3 Select Pushdown com o PrestoDB ou Trino

Para habilitar o S3 Select Pushdown para Presto no Amazon EMR, use a classificação de configuração `presto-connector-hive` para definir `hive.s3select-pushdown.enabled` como `true`, conforme mostrado no exemplo a seguir. Para obter mais informações, consulte Configurar aplicações. O valor hive.s3select-pushdown.max-connections também deve ser definido. Para a maioria dos aplicativos, a configuração padrão de *500* deve ser adequada. Para obter mais informações, consulte Noções básicas e ajuste do hive.s3select-pushdown.max-connections abaixo.

No PrestoSQL nas versões 6.1.0 a 6.3.0 do EMR, substitua *presto-connector-hive* no exemplo abaixo por `prestosql-connector-hive`.

As versões 6.4.0 e posteriores do Amazon EMR usam o novo nome Trino em vez de PrestoSQL. Se você usa o Trino, substitua *presto-connector-hive* no exemplo abaixo por `trino-connector-hive`

```
[
    {
        "classification": "presto-connector-hive",
        "properties": {
            "hive.s3select-pushdown.enabled": "true",
            "hive.s3select-pushdown.max-connections": "500"
        }
    }
]
```

Noções básicas e ajuste do hive.s3select-pushdown.max-connections

Por padrão, o Presto usa o EMRFS como seu sistema de arquivos. A configuração `fs.s3.maxConnections` na classificação de configuração `emrfs-site` especifica o máximo permitido de conexões de cliente com o Amazon S3 por meio do EMRFS para Presto. Por padrão, é 500. O S3 Select Pushdown ignora o EMRFS ao acessar o Amazon S3 para operações de predicado. Nesse caso, o valor de `hive.s3select-pushdown.max-connections` determina o número máximo de conexões de cliente permitidas para essas operações de nós do operador. No entanto, quaisquer solicitações para o Amazon S3 que o Presto inicia que não são enviadas, como operações GET, continuam a ser regidas pelo valor de `fs.s3.maxConnections`.

Se o seu aplicativo tiver o erro "Tempo limite atingido em espera para conexão do grupo", aumente o valor de `hive.s3select-pushdown.max-connections` e `fs.s3.maxConnections`.

Adicionar conectores de banco de dados

Você pode usar classificações de configuração para configurar as propriedades do conector JDBC ao criar um cluster. As classificações de configuração começam com `presto-connector`, por exemplo, `presto-connector-postgresql`. As classificações de configuração disponíveis dependem da versão do Amazon EMR. Para obter as classificações de configuração disponíveis com a versão mais recente, consulte the section called “Classificações de configuração 5.36.2” para Amazon EMR 5.36.2. Se você estiver usando uma versão diferente do Amazon EMR, consulte Versões de lançamento 5.x do Amazon EMR para conhecer as classificações de configuração. Para obter mais informações sobre as propriedades que podem ser configuradas com cada conector, consulte https://prestodb.io/docs/current/connector.html.

Example — configurar um cluster com o conector JDBC do PostgreSQL

Para iniciar um cluster com o conector do PostgreSQL instalado e configurado, crie um arquivo JSON que especifique a classificação de configuração, por exemplo, `myConfig.json` com o seguinte conteúdo e salve-o localmente.

Substitua as propriedades da conexão, conforme apropriado para sua configuração e como mostrado no tópico Conectos do PostgreSQL na documentação do Presto.

```
[
  {
    "Classification": "presto-connector-postgresql",
```

```
    "Properties": {
      "connection-url": "jdbc:postgresql://example.net:5432/database",
      "connection-user": "MYUSER",
      "connection-password": "MYPASS"
    },
    "Configurations": []
  }
]
```

Quando você cria o cluster, referencie o caminho para o arquivo JSON usando a opção `--configurations`, conforme mostrado no exemplo a seguir, em que `myConfig.json` está no mesmo diretório onde você executa o comando:

```
aws emr create-cluster --name PrestoConnector --release-label emr-5.36.2 --instance-
type m5.xlarge \
--instance-count 2 --applications Name=Hadoop Name=Hive Name=Pig Name=Presto \
--use-default-roles --ec2-attributes KeyName=myKey \
--log-uri s3://amzn-s3-demo-bucket/logs --enable-debugging \
--configurations file://myConfig.json
```

Usar SSL/TLS e configurar LDAPS com o Presto no Amazon EMR

Com as versões 5.6.0 e posteriores do Amazon EMR, você pode habilitar SSL/TLS para ajudar a proteger a comunicação interna entre nós do Presto. Você pode fazer isso definindo uma configuração de segurança para criptografia em trânsito. Para obter mais informações, consulte Opções de criptografia e Usar configurações de segurança para definir a segurança do cluster no Guia de gerenciamento do Amazon EMR.

Quando você usa uma configuração de segurança com criptografia em trânsito, o Amazon EMR faz o seguinte para o Presto:

- Distribui os artefatos ou certificados de criptografia que você especifica para a criptografia em trânsito em todo o cluster do Presto. Para obter mais informações, consulte Fornecer certificados para criptografia de dados em trânsito.
- Define as seguintes propriedades usando a classificação de configuração `presto-config`, que corresponde ao arquivo `config.properties` para o Presto:
  - Define `http-server.http.enabled` como `false` em todos os nós, o que desabilita HTTP em favor de HTTPS. Isso exige que você forneça certificados que funcionem para DNS público e privado quando definir a configuração de segurança para criptografia em trânsito. Uma maneira

de fazer isso é usar certificados SAN (Subject Alternative Name), compatíveis com vários domínios.

- Define os valores `http-server.https.*`. Para obter detalhes da configuração, consulte Autenticação LDAP na documentação do Presto.

Além disso, com as versões 5.10.0 e posteriores do Amazon EMR, você pode configurar a autenticação LDAP para conexões de clientes ao coordenador do Presto usando HTTPS. Essa configuração usa um LDAP seguro (LDAPS). O TLS deve ser habilitado no servidor LDAP, e o cluster do Presto deve usar uma configuração de segurança com a criptografia de dados em trânsito habilitada. Configurações adicionais são necessárias. As opções de configuração são diferentes de acordo com a versão do Amazon EMR que você usa. Para obter mais informações, consulte Usar autenticação LDAP para o Presto no Amazon EMR.

O Presto no Amazon EMR usa a porta 8446 para HTTPS interno por padrão. A porta usada para comunicação interna deve ser a mesma usada para o acesso HTTPS de clientes ao coordenador do Presto. A propriedade `http-server.https.port` na classificação de configuração `presto-config` especifica a porta.

Usar autenticação LDAP para o Presto no Amazon EMR

Siga as etapas nesta seção para configurar LDAP. Consulte cada etapa para obter exemplos e links para mais informações.

Etapas para configurar a autenticação LDAP

- Etapa 1: coletar informações sobre o servidor LDAP e copiar o certificado do servidor para o Amazon S3
- Etapa 2: definir uma configuração de segurança
- Etapa 3: criar uma configuração JSON com propriedades do Presto para LDAP
- Etapa 4: criar o script para copiar o certificado do servidor LDAP e carregá-lo no Amazon S3
- Etapa 5: criar um cluster

Etapa 1: coletar informações sobre o servidor LDAP e copiar o certificado do servidor para o Amazon S3

Você precisará dos itens e informações na seção a seguir do seu servidor LDAP para configurar a autenticação LDAP.

O endereço IP ou o nome do host do servidor LDAP

O coordenador do Presto no nó principal do Amazon EMR deve ser capaz de acessar o servidor LDAP no endereço IP ou nome de host especificado. Por padrão, o Presto se comunica com o servidor LDAP usando LDAPS pela porta 636. Se sua implementação do LDAP exige uma porta personalizada, você pode especificá-la usando a propriedade `ldap.url` com as versões 5.16.0 ou posteriores do Amazon EMR ou usando `authentication.ldap.url` com versões anteriores. Substitua a porta personalizada em `636`, conforme mostrado nos exemplos da classificação de configuração `presto-config` em Etapa 3: criar uma configuração JSON com propriedades do Presto para LDAP. Verifique se todos os firewalls e grupos de segurança permitem o tráfego de entrada e saída na porta 636 (ou na porta personalizada) e também na porta 8446 (ou na porta personalizada), que é usada para comunicações do cluster interno.

O certificado do servidor LDAP

Você deve carregar o arquivo do certificado em um local seguro no Amazon S3. Para obter mais informações, consulte Como carregar arquivos e pastas em um bucket do S3 no Guia do usuário do Amazon Simple Storage Service. Crie uma ação de bootstrap que copie esse certificado do Amazon S3 em cada nó do cluster quando esse cluster é iniciado. Em Etapa 4: criar o script para copiar o certificado do servidor LDAP e carregá-lo no Amazon S3. O certificado de exemplo é *`s3://amzn-s3-demo-bucket/ldap_server.crt`*.

As configurações do servidor LDAP para associação anônima

Se a associação anônima estiver desabilitada no PrestoDB, você precisará do ID de usuário (UID) e da senha de uma conta com permissões para associação ao servidor LDAP, para que o servidor PrestoDB possa estabelecer uma conexão. Especifique o UID e a senha usando as propriedades `internal-communication.authentication.ldap.user` e `internal-communication.authentication.ldap.password` na classificação de configuração `presto-config`. O Amazon EMR 5.10.0 não é compatível com essas configurações, portanto, a associação anônima deve ser compatível com o servidor LDAP quando você usa esta versão.

Observe que o Trino não exige a configuração de vinculação anônima.

Para obter o status da associação anônima no servidor LDAP

- Use o comando ldapwhoami do cliente Linux, conforme mostrado no exemplo a seguir:

```
ldapwhoami -x -H ldaps://LDAPServerHostNameOrIPAddress
```

Se a associação anônima não for permitida, o comando retornará o seguinte:

```
ldap_bind: Inappropriate authentication (48)
additional info: anonymous bind disallowed
```

Para verificar se uma conta tem as permissões para um servidor LDAP que usa autenticação simples

- Use o comando ldapwhoami do cliente Linux, conforme mostrado no exemplo a seguir. O exemplo usa um usuário fictício,*presto*, armazenado em um servidor Open LDAP executado em uma EC2 instância com o nome de host fictício. *ip-xxx-xxx-xxx-xxx.ec2.internal* O usuário está associado à unidade organizacional (OU) *admins* e à senha*123456*:

  ```
  ldapwhoami -x -w "123456" -D uid=presto,ou=admins,dc=ec2,dc=internal -H ldaps://ip-xxx-xxx-xxx-xxx.ec2.internal
  ```

  Se a conta for válida e tiver as permissões adequadas, o comando retornará:

  ```
  dn:uid=presto,ou=admins,dc=ec2,dc=internal
  ```

As configurações de exemplo em Etapa 3: criar uma configuração JSON com propriedades do Presto para LDAP incluem essa conta para maior clareza, com exceção do exemplo de 5.10.0, em que ela não é compatível. Se o servidor LDAP usa a associação anônima, remova `internal-communication.authentication.ldap.user` e os pares de nome/valor `internal-communication.authentication.ldap.password`.

O nome diferenciado (DN) LDAP para usuários do Presto

Ao especificar a configuração LDAP para o Presto, você especifica um padrão de associação que consiste `${USER}` em uma unidade organizacional (OU) e componentes de domínio adicionais (). DCs O Presto substitui `${USER}` pelo ID de usuário (UID) real de cada usuário durante a autenticação de senha para corresponder ao nome diferenciado (DN) que esse padrão de associação especifica. Você precisa do OUs que os usuários qualificados pertencem e deles DCs. Por exemplo, para permitir que os usuários da OU `admins` no domínio `corp.example.com` autentiquem para o Presto, especifique `${USER},ou=admins,dc=corp,dc=example,dc=com` como o padrão de associação do usuário.

> **Note**
>
> Quando você usa AWS CloudFormation, precisa usar a função Fn: :Sub para substituí-la `${USER}` pelo ID de usuário (UID) real. Para obter mais informações, consulte o tópico Fn::Sub no Guia do usuário do AWS CloudFormation .

Ao usar o Amazon EMR 5.10.0, você só pode especificar um padrão desse tipo. Ao usar as versões 5.11.0 ou posteriores do Amazon EMR, você pode especificar vários padrões separados por dois-pontos (:). Os usuários que tentam autenticar para o Presto são comparados com o primeiro padrão, o segundo, e assim por diante. Para obter um exemplo, consulte Etapa 3: criar uma configuração JSON com propriedades do Presto para LDAP.

Etapa 2: definir uma configuração de segurança

Crie uma configuração de segurança com a criptografia em trânsito habilitada. Para obter mais informações, consulte Criar uma configuração de segurança no Guia de gerenciamento do Amazon EMR. Os artefatos de criptografia que você fornece ao configurar a criptografia em trânsito são usados para criptografar a comunicação interna entre os nós Presto. Para obter mais informações, consulte Fornecer certificados para criptografia de dados em trânsito. O certificado do servidor LDAP é usado para autenticar conexões de clientes ao servidor Presto.

Etapa 3: criar uma configuração JSON com propriedades do Presto para LDAP

Você pode usar a classificação de configuração `presto-config` para definir as propriedades do Presto para LDAP. O formato e o conteúdo de `presto-config` são ligeiramente diferentes, dependendo da versão do Amazon EMR e da instalação do Presto (PrestoDB ou Trino). Exemplos de diferenças de configuração serão fornecidos posteriormente nesta seção. Para obter mais informações, consulte Configurar aplicações.

As etapas a seguir pressupõem que você salve os dados do JSON em um arquivo. *`MyPrestoConfig.json`* Se você usa o console, carregue o arquivo em um local seguro no Amazon S3 para poder fazer referência a ele quando criar o cluster. Se você usar o AWS CLI, poderá referenciar o arquivo localmente.

Example Versões 6.1.0 e posteriores do Amazon EMR com PrestoSQL (Trino)

O exemplo a seguir usa o nome de host LDAP de Etapa 1: coletar informações sobre o servidor LDAP e copiar o certificado do servidor para o Amazon S3 para autenticação no servidor LDAP para

associação. Dois padrões de associação do usuário são especificados, o que indica que os usuários dentro da unidade organizacional (UO) `admins` e da UO `datascientists` no servidor LDAP estão qualificados para autenticação ao servidor Trino como usuários. Os padrões de associação são separados com uma vírgula (`:`).

As versões 6.4.0 e posteriores do Amazon EMR usam o novo nome Trino em vez de PrestoSQL. Se você usa o Trino, substitua *`prestosql-config`* na seguinte classificação de configuração por `trino-config` e *`prestosql-password-authenticator`* por `trino-password-authenticator`.

```
[
   {
      "Classification":"prestosql-config",
      "Properties":{
         "http-server.authentication.type":"PASSWORD"
      }
   },
   {
      "Classification":"prestosql-password-authenticator",
      "Properties":{
         "password-authenticator.name":"ldap",
         "ldap.url":"ldaps://ip-xxx-xxx-xxx-xxx.ec2.internal:636",
         "ldap.user-bind-pattern": "uid=${USER},ou=admins,dc=ec2,dc=internal:uid=
${USER},ou=datascientists,dc=ec2,dc=internal"
      }
   }
]
```

Example Versões 5.16.0 e posteriores do Amazon EMR

O exemplo a seguir usa o ID de usuário, a senha e o nome de host LDAP de Etapa 1: coletar informações sobre o servidor LDAP e copiar o certificado do servidor para o Amazon S3 para autenticar ao servidor LDAP para associação. Dois padrões de associação do usuário são especificados, o que indica que os usuários dentro da OU `admins` e da OU `datascientists` no servidor LDAP estão qualificados para autenticação ao servidor Presto como usuários. Os padrões de associação são separados com uma vírgula (`:`).

```
[{
        "Classification": "presto-config",
                "Properties": {
                        "http-server.authentication.type": "PASSWORD"
```

```
                }
        },
        {
                "Classification": "presto-password-authenticator",
                "Properties": {
                        "password-authenticator.name": "ldap",
                        "ldap.url": "ldaps://ip-xxx-xxx-xxx-xxx.ec2.internal:636",
                        "ldap.user-bind-pattern": "uid=
${USER},ou=admins,dc=ec2,dc=internal:uid=${USER},ou=datascientists,dc=ec2,dc=internal",
                        "internal-communication.authentication.ldap.user": "presto",
                        "internal-communication.authentication.ldap.password": "123456"
                }
        }]
```

Example Amazon EMR 5.11.0 a 5.15.0

O formato da classificação de configuração `presto-config` é um pouco diferente para essas versões. O exemplo a seguir especifica os mesmos parâmetros que o exemplo anterior.

```
[{
        "Classification": "presto-config",
                "Properties": {
                        "http-server.authentication.type": "LDAP",
                        "authentication.ldap.url": "ldaps://ip-xxx-xxx-xxx-
xxx.ec2.internal:636",
                        "authentication.ldap.user-bind-pattern": "uid=
${USER},ou=admins,dc=ec2,dc=internal:uid=${USER},ou=datascientists,dc=ec2,dc=internal",
                        "internal-communication.authentication.ldap.user": "presto",
                        "internal-communication.authentication.ldap.password": "123456"
                }
        }]
```

Example Amazon EMR 5.10.0

O Amazon EMR 5.10.0 só é compatível com associação anônima. Portanto, essas entradas são omitidas. Além disso, apenas um único padrão de associação pode ser especificado.

```
[{
        "Classification": "presto-config",
                "Properties": {
                        "http-server.authentication.type": "LDAP",
                        "authentication.ldap.url": "ldaps://ip-xxx-xxx-xxx-
xxx.ec2.internal:636",
```

```
                        "ldap.user-bind-pattern": "uid=
${USER},ou=prestousers,dc=ec2,dc=internal"
                }
        }]
```

Etapa 4: criar o script para copiar o certificado do servidor LDAP e carregá-lo no Amazon S3

Crie um script que copie o arquivo do certificado em cada dó no cluster e o adicione ao repositório de chaves. Crie o script usando um editor de texto, salve-o e carregue-o no Amazon S3. EmEtapa 5: criar um cluster, o arquivo de script é referenciado como*s3://amzn-s3-demo-bucket/LoadLDAPCert.sh*.

O script de exemplo a seguir usa a senha padrão do keystore,*changeit*. Recomendamos que você se conecte ao nó principal após criar o cluster e alterar a senha do repositório de chaves usando o comando keytool.

```
#!/bin/bash
aws s3 cp s3://amzn-s3-demo-bucket/ldap_server.crt .
sudo keytool -import -keystore /usr/lib/jvm/jre-1.8.0-openjdk.x86_64/lib/security/
cacerts -trustcacerts -alias ldap_server -file ./ldap_server.crt -storepass changeit -
noprompt
```

Etapa 5: criar um cluster

Ao criar o cluster, especifique o Presto e outras aplicações que você deseja que o Amazon EMR instale. Os exemplos a seguir também fazem referência às propriedades da classificação de configuração em um JSON, mas você também pode especificar a classificação de configuração em linha.

Criar um cluster do Presto com autenticação LDAP usando o console do Amazon EMR

1. Navegue até o novo console do Amazon EMR e selecione Alternar para o console antigo na navegação lateral. Para obter mais informações sobre o que esperar ao alternar para o console antigo, consulte Usar o console antigo.
2. Escolha Create cluster (Criar cluster), Go to advanced options (Ir para opções avançadas).
3. Selecione Presto e outras aplicações para o Amazon EMR instalar e, em Configuração de software, selecione a Versão do Amazon EMR a ser usada. A autenticação LDAP só é compatível com as versões 5.10.0 e posteriores do Amazon EMR.

4. Em Editar configurações de software, selecione Carregar JSON do S3, insira o local no Amazon S3 do arquivo de configuração JSON que você criou em Etapa 3: criar uma configuração JSON com propriedades do Presto para LDAP e, em seguida, selecione Próximo.
5. Configure os recursos de hardware e rede do cluster e, em seguida, selecione Next (Próximo).
6. Selecione Bootstrap Actions (Ações de bootstrap). Em Add bootstrap action (Adicionar ação de bootstrap), selecione Custom action (Personalizar ação) e, em seguida, selecione Configure and add (Configurar e adicionar).
7. Insira um nome para a ação de bootstrap, insira a localização do script que você criouEtapa 4: criar o script para copiar o certificado do servidor LDAP e carregá-lo no Amazon S3, por exemplo, s3://amzn-s3-demo-bucket/Load LDAPCert .sh, e escolha Adicionar.
8. Em General Options (Opções gerais), Tags e Additional Options (Opções adicionais), selecione as configurações adequadas para o seu aplicativo e, em seguida, selecione Next (Próximo).
9. Selecione Authentication and encryption (Autenticação e criptografia) e, em seguida, selecione a Security configuration (Configuração de segurança) que você criou em Etapa 2: definir uma configuração de segurança.
10. Selecione outras opções de segurança conforme adequado para o seu aplicativo e, em seguida, selecione Create cluster (Criar cluster).

Para criar um cluster do Presto com autenticação LDAP usando a AWS CLI

- Use o comando `aws emr create-cluster`. No mínimo, especifique o aplicativo do Presto e também a classificação de configuração, o script de bootstrap e a configuração de segurança do Presto que você criou nas etapas anteriores. O exemplo a seguir faz referência ao arquivo de configuração como um arquivo JSON salvo no mesmo diretório em que você executa o comando. Por outro lado, o script de bootstrap deve ser salvo no Amazon S3. O exemplo a seguir usa `s3://amzn-s3-demo-bucket/LoadLDAPCert.sh`.

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr create-cluster  --applications Name=presto --release-label emr-5.16.0 \
```

```
--use-default-roles --ec2-attributes KeyName=MyKeyPair,SubnetId=subnet-1234ab5
 \ --instance-count 3 --instance-type m5.xlarge --region us-west-2 --name
 "MyPrestoWithLDAPAuth" \
--bootstrap-actions Name="Distribute LDAP server cert",Path="s3://amzn-s3-demo-
bucket/LoadLDAPCert.sh" \
--security-configuration MyPrestoLDAPSecCfg --configurations file://
MyPrestoConfig.json
```

Ativar o modo estrito do Presto

Em determinadas situações, consultas de execução prolongada podem gerar altos custos e fazer com que o Amazon EMR use mais recursos de cluster. Isso retira recursos de outras workloads do cluster. Com as versões 6.8 e posteriores do Amazon EMR, você pode usar um atributo de modo estrito que rejeita ou avisa sobre os seguintes tipos de consultas de execução prolongada:

- Consultas sem predicados nas colunas particionadas que resultam em verificações de tabelas de grandes quantidades de dados
- Consultas com junções cruzadas entre duas tabelas grandes
- Consultas que classificam um grande número de linhas sem limite

Depois que o Presto otimiza por completo o plano de consulta, o modo estrito é executado. Para usar e personalizar o modo estrito de acordo com suas necessidades de consulta, configure o Presto das maneiras a seguir.

Configurações do Presto para modo estrito

Configuração	Descrição	Padrão
`strict-mode-enabled`	Ativa e desativa o modo estrito. Um valor de `true` indica que o modo estrito está ativado.	`false`
`strict-mode-fail-query`	Rejeita consultas se o modo estrito detecta prováveis consultas de execução prolongada. No caso de	`false`

Configuração	Descrição	Padrão
	`false`, o Amazon EMR só emite um aviso.	
`strict-mode-restrictions`	Especifica as restrições a serem aplicadas quando o modo estrito está ativado. O modo estrito é compatível com as seguintes restrições: MANDATORY_PARTITION_PREDICATE, DISALLOW_CROSS_JOIN e LIMITED_SORT.	MANDATORY_PARTITION_PREDICATE,DISALLOW_CROSS_JOIN, LIMITED_SORT

Para testar o modo estrito, você pode substituir essas configurações ou defini-las como propriedades da sessão ao usar o cliente do Presto.

Definir a configuração na criação do cluster com a AWS Management Console

1. Escolha Criar cluster e selecione Amazon EMR versão 6.8.0 e Presto ou Trino. Para obter mais informações, consulte Instalar PrestoDB e Trino.
2. Especifique as propriedades de configuração para o modo estrito diretamente ou carregue um arquivo JSON para o Amazon S3. Como alternativa, selecione o AWS Glue Data Catalog como metastore. Especifique a VPC, as sub-redes, as ações de bootstrap, o par de chaves e o grupo de segurança. Para criar o cluster, escolha Criar cluster.
3. Faça login no nó primário do cluster e execute `presto-cli` ou `trino-cli`.
4. Envie suas consultas. O modo estrito valida cada consulta e determina se ela é de execução prolongada. Dependendo da configuração de `strict-mode-fail-query`, o Amazon EMR rejeita a consulta ou emite um aviso.
5. Ao concluir as consultas, encerre o cluster e exclua seus recursos.

Definir a configuração em um cluster em execução com a AWS CLI

1. Faça login no nó primário do seu cluster com o AWS CLI e execute `presto-cli` outrino-cli.

2. Execute os comandos a seguir com os valores desejados.

```
set session strict_mode_enabled = true;
set session strict_mode_fail_query = false;
set session strict_mode_restrictions = 'DISALLOW_CROSS_JOIN,LIMITED_SORT';
```

Considerações

Quando usar o modo estrito, considere os seguintes pontos:

- Em alguns casos, o modo estrito pode rejeitar consultas de execução curta que não consomem muitos recursos. Por exemplo, consultas em tabelas pequenas não aplicam filtragem dinâmica nem substituem junções internas por junções cruzadas. Isso pode fazer com que a consulta use o predicado de partição obrigatória ou proíba junção cruzada. Quando isso acontece, o modo estrito rejeita a consulta.
- A verificação do modo estrito só é aplicada nos tipos de consulta SELECT, INSERT, CREATE TABLE AS SELECT e EXPLAIN ANALYZE.
- Você só pode usar a restrição no predicado de partição obrigatória com o conector do Hive.

Lidar com a perda de instâncias spot no Presto

Com instâncias spot no Amazon EMR, você pode executar workloads de big data na EC2 capacidade não utilizada da Amazon a um custo reduzido. Em troca por um custo mais baixo, a Amazon EC2 pode interromper instâncias spot com uma notificação de dois minutos. Quando você encerra um nó, o Presto pode levar até 10 minutos para retornar um erro. Isso causa atrasos desnecessários nos relatórios de erros e possíveis novas tentativas. O encerramento rápido é um atributo que permite o controle da forma como o Presto manipula nós encerrados.

O trabalho do coordenador do Presto é acompanhar todos os nós de processamento com sondagens regulares dos seus status. Sem um encerramento rápido, o coordenador não consulta o YARN NodeManager para conhecer o status de cada nó. Isso pode resultar em um longo loop de novas tentativas antes que a consulta apresente falha. Com o encerramento rápido, o coordenador do Presto consulta o status do nó NodeManager assim que a sondagem deixa de acessar o host. Se NodeManager mostra que o nó está inativo, o Presto abandona novas tentativas, apresenta falha na consulta e retorna um erro. `NODE_DECOMMISSIONED`

O conjunto de parâmetros de configuração a seguir permite o controle e a personalização do comportamento do Presto no caso do encerramento do nó.

Configurações do Presto para tratamento de falhas em nós

Configuração	Descrição	Padrão
`query.remote-task.max-backoff-duration`	Período durante o qual o coordenador continua tentando obter o status da tarefa remota nos nós de processamento.	10 minutos
`query.remote-task.quick-terminate-node-failure`	Ativa a falha rápida do nó se o coordenador não consegue acessar o nó ou não consegue se conectar ao worker executado nesse nó. O valor de `query.remote-task.terminate-on-connect-exception` determina se o coordenador deve acessar o nó ou se conectar ao worker. O nó apresenta falha na consulta e o Amazon EMR remove o nó da lista de workers disponíveis. Quando isso acontece, você não pode usar o nó para programar novas consultas. Quando você define esse valor como `false`, o Presto volta ao comportamento anterior, em que o coordenador do Presto novamente tenta acessar o nó (para	`true`

Configuração	Descrição	Padrão
	`query.remote-task.max-backoff-duration`) antes de marcar o nó como indisponível e indica falha na consulta em andamento no nó.	
`query.remote-task.terminate-on-connect-exception`	Especifica se o Amazon EMR deve ser um nó se o host está acessível, mas o coordenador não consegue se conectar ao processo do worker do host. Ao definir esse valor como `true`, você ativa falha na consulta rápida se o host está inacessível.	`false`

Usar a escalabilidade automática do Presto com desativação tranquila

As versões 5.30.0 e posteriores do Amazon EMR incluem um atributo que você pode usar para definir um período de carência para determinadas ações de escalabilidade. O período de carência possibilita que as tarefas do Presto continuem em execução antes que o nó seja encerrado devido a uma ação de redimensionamento para reduzir a escala horizontalmente ou a uma solicitação de política de escalabilidade automática. Para obter mais informações sobre regras de escalabilidade, consulte Noções básicas das regras de escalabilidade automática no Guia de gerenciamento do Amazon EMR. A escalabilidade automática do Presto com desativação tranquila impede que novas tarefas sejam programadas em um nó que está sendo desativado e, ao mesmo tempo, permite que as tarefas que já estão sendo executadas sejam concluídas antes que o tempo limite de desligamento seja atingido. As consultas em execução concluirão a execução antes que o nó seja desativado. A escalabilidade automática não é compatível com frotas de instâncias.

Você pode controlar quanto tempo deve ser permitido até a conclusão das tarefas do Presto após o recebimento de uma solicitação de desligamento de escalabilidade automática. Por padrão, o tempo

limite de desligamento do Amazon EMR é `0` minuto, o que significa que o Amazon EMR encerra imediatamente o nó e todas as tarefas do Presto executadas nele, se exigido por uma solicitação para reduzir a escala horizontalmente.. Para definir um tempo limite mais longo para as tarefas do Presto no Amazon EMR, permitindo que as consultas em execução sejam concluídas antes da ação de reduzir a escala verticalmente de um cluster, use a classificação de configuração `presto-config` para definir o parâmetro `graceful-shutdown-timeout` como um valor em segundos ou minutos maior que zero. Para obter mais informações, consulte Configurar aplicações.

Por exemplo, aumentar o valor de `graceful-shutdown-timeout` para `"30m"` especifica um tempo limite de 30 minutos. Após o término do tempo limite de desligamento, o nó marcado para desativação será encerrado de maneira forçada se estiver aguardando a conclusão das tarefas de consulta e a consulta apresentará falha. Se as tarefas de consulta forem concluídas em cinco minutos, o nó marcado para desativação será encerrado em cinco minutos, desde que outras aplicações do YARN tenham concluído a execução.

Example Exemplo de configuração de escalabilidade automática do Presto com desativação tranquila

Substitua o valor `graceful-shutdown-timeout` pelo número de minutos apropriado para a configuração. Não há um valor máximo. O exemplo abaixo define um valor de tempo limite de `1800` segundos (30 minutos).

```
[
    {
        "classification": "presto-config",
        "properties": {
            "graceful-shutdown-timeout": "1800s"
        }
    }
]
```

Limitações

A desativação tranquila do PrestoDB não funciona em clusters do EMR em que a conectividade HTTP está desabilitada, como quando `http-server.http.enabled` está definido como `false`. O Trino não é compatível com a desativação tranquila, independentemente da configuração de `http-server.http.enabled`.

Considerações com o Presto no Amazon EMR

Considere as limitações a seguir quando executar o Presto no Amazon EMR.

Executável da linha de comando do Presto

No Amazon EMR, o PrestoDB e o Trino usam o mesmo executável da linha de comando, `presto-cli`, como no exemplo a seguir.

```
presto-cli --catalog hive
```

Propriedades de implantação do Presto não configuráveis

A versão do Amazon EMR que você usa determina as configurações de implantação do Presto que estão disponíveis. Para obter mais informações sobre essas propriedades de configuração, consulte Implantar o Presto na documentação do Presto. A tabela a seguir mostra as diferentes opções de configuração para arquivos de `properties` do Presto.

Arquivo	Configurável
`log.properties`	PrestoDB: configurável nas versões 4.0.0 e posteriores do Amazon EMR. Use a classific ação para a configuração `presto-log` .
`config.properties`	PrestoDB: configurável nas versões 4.0.0 e posteriores do Amazon EMR. Use a classific ação para a configuração `presto-config` .
`hive.properties`	PrestoDB: configurável nas versões 4.1.0 e posteriores do Amazon EMR. Use a classific ação para a configuração `presto-co nnector-hive` .
`node.properties`	PrestoDB: configurável nas versões 5.6.0 e posteriores do Amazon EMR. Use a classific ação para a configuração `presto-node` .
`jvm.config`	Não configurável.

Instalação do PrestoDB

O nome da aplicação, Presto, continua a ser usado para instalar o PrestoDB em clusters.

Você pode instalar o PrestoDB ou o Trino, mas não pode instalar os dois em um único cluster. Se o PrestoDB e o Trino forem especificados na tentativa de criação de um cluster, ocorrerá um erro de validação e a solicitação de criação do cluster falhará.

Configuração do EMRFS e do PrestoS3 FileSystem

Com as versões 5.12.0 e posteriores do Amazon EMR, o PrestoDB pode usar o EMRFS. Para obter mais informações, consulte EMR File System (EMRFS) no Guia de gerenciamento do Amazon EMR. Nas versões anteriores do Amazon EMR, o PrestoS3 FileSystem é a única opção de configuração.

Você pode usar uma configuração de segurança para configurar a criptografia para dados do EMRFS no Amazon S3. Você também pode usar perfis do IAM para solicitações do EMRFS para o Amazon S3. Para obter mais informações, consulte Noções básicas sobre opções de criptografia e Configurar perfis do IAM para solicitações do EMRFS para o Amazon S3 no Guia de gerenciamento do Amazon EMR.

> **Note**
>
> Se você consultar dados subjacentes no Amazon S3 com a versão 5.12.0 do Amazon EMR, poderão ocorrer erros no Presto. Isso acontece porque o Presto não consegue obter valores de classificação de configuração em `emrfs-site.xml`. Como solução alternativa, crie um subdiretório `emrfs` em `usr/lib/presto/plugin/hive-hadoop2/` e crie um link simbólico em `usr/lib/presto/plugin/hive-hadoop2/emrfs` para o arquivo `/usr/share/aws/emr/emrfs/conf/emrfs-site.xml` existente. Em seguida, reinicie o processo presto-server (`sudo presto-server stop` seguido por `sudo presto-server start`).

Você pode substituir o padrão do EMRFS e usar o PrestOS3 em vez disso. FileSystem Para fazer isso, use a classificação de configuração `presto-connector-hive` para definir `hive.s3-file-system-type` como `PRESTO` , conforme mostrado no exemplo a seguir. Para obter mais informações, consulte Configurar aplicações.

```
[
   {
```

```
      "Classification": "presto-connector-hive",
      "Properties": {
         "hive.s3-file-system-type": "PRESTO"
      }
   }
]
```

Se você usa o PrestOS3FileSystem, use a classificação de `presto-connector-hive` configuração para configurar as propriedades do FileSystem PrestOS3. Para obter mais informações sobre as propriedades disponíveis, consulte Configuração do Amazon S3 na seção Conector do Hive da documentação do Presto. Estas configurações não se aplicam ao EMRFS.

Configuração padrão para personificação do usuário final

Por padrão, as versões 5.12.0 e posteriores do Amazon EMR possibilitam a personificação do usuário final para o acesso ao HDFS. Para obter mais informações, consulte Personificação de usuário final na documentação do Presto. Para alterar essa configuração usando a classificação de configuração `presto-config`, defina a propriedade `hive.hdfs.impersonation.enabled` como `false`.

Porta padrão para interface Web do Presto

Por padrão, o Amazon EMR configura a interface Web do Presto no coordenador do Presto para usar a porta 8889 (para PrestoDB e Trino). Para alterar a porta, use a classificação de configuração `presto-config` para definir a propriedade `http-server.http.port`. Para obter mais informações, consulte Propriedades de configuração na seção Implantar o Presto na documentação do Presto.

Problema com a execução do bucket do Hive em algumas versões

O Presto versão 152.3 tem um problema com a execução do bucket do Hive que prejudica significativamente o desempenho das consultas do Presto em algumas circunstâncias. As versões 5.0.3, 5.1.0 e 5.2.0 do Amazon EMR incluem essa versão do Presto. Para reduzir o problema, use a classificação de configuração `presto-connector-hive` para definir a propriedade `hive.bucket-execution` como `false`, conforme mostrado no exemplo a seguir.

```
[
   {
      "Classification": "presto-connector-hive",
      "Properties": {
```

```
        "hive.bucket-execution": "false"
      }
   }
]
```

Histórico de versões do Presto

A tabela a seguir informa a versão do Presto incluída em cada versão do Amazon EMR e os componentes instalados com a aplicação. Para obter as versões dos componentes em cada versão, consulte a seção Versão de componentes da sua versão em Versões 7.x do Amazon EMR, Versões de lançamento 6.x do Amazon EMR ou Versões de lançamento 5.x do Amazon EMR.

Informações de versão do Presto

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
amzn-7.9.0	0,287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
amzn-7.8.0	0,287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
		client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-7.7.0	0,287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-7.6.0	0,287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-7.5.0	0,287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-7.4.0	0,287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-7.3.0	0,285	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-7.2.0	0,285	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.36.2	0,267	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-7.1.0	0,284	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-7.0.0	0,283	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.15.0	0,283	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-6.14.0	0,281	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.13.0	0,281	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-6.12.0	0,281	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.11.1	0,279	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-6.11.0	0,279	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.10.1	0,278	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-6.10.0	0,278	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.9.1	0.276	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-6.9.0	0.276	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.8.1	0,273	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-6.8.0	0,273	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.7.0	0,272	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.36.1	0,267	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.36.0	0,267	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-6.6.0	0,267	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.35.0	0,266	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-6.5.0	0,261	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.4.0	0.254.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-6.3.1	0.245.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.3.0	0.245.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-6.2.1	0.238.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.2.0	0.238.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-6.1.1	0.232	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.1.0	0.232	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-6.0.1	0.230	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.0.0	0.230	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.34.0	0,261	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.33.1	0.245.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.33.0	0.245.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.32.1	0.240.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.32.0	0.240.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.31.1	0.238.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.31.0	0.238.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.30.2	0.232	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.30.1	0.232	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.30.0	0.232	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.29.0	0.227	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.28.1	0.227	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.28.0	0.227	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.27.1	0.224	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.27.0	0.224	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.26.0	0.220	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.25.0	0.220	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.24.1	0.219	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.24.0	0.219	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.23.1	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.23.0	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.22.0	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.21.2	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.21.1	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.21.0	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.20.1	0.214	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.20.0	0.214	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.19.1	0.212	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.19.0	0.212	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.18.1	0.210	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.18.0	0.210	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.17.2	0.206	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.17.1	0.206	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.17.0	0.206	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.16.1	0.203	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.16.0	0.203	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.15.1	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.15.0	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.14.2	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.14.1	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.14.0	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.13.1	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.13.0	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.12.3	0.188	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.12.2	0.188	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.12.1	0.188	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.12.0	0.188	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.11.4	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.11.3	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.11.2	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.11.1	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.11.0	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.10.1	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.10.0	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.9.1	0.184	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.9.0	0.184	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.8.3	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.8.2	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.8.1	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.8.0	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.7.1	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.7.0	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.6.1	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.6.0	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.5.4	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.5.3	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.5.2	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.5.1	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.5.0	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.4.1	0.166	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.4.0	0.166	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.3.2	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.3.1	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.3.0	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.2.3	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.2.2	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.2.1	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.2.0	0.152.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.1.1	0.152.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.1.0	0.152.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.0.3	0.152.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.0.2	0.150	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.0.1	0.150	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Rótulo de versão do Amazon EMR	Versão do Presto	Componentes instalados com o Presto
emr-5.0.0	0.150	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Apache Spark

O Apache Spark é um modelo de framework de processamento e de programação distribuído que ajuda você a executar machine learning, processamento de streams ou analytics de gráficos com clusters do Amazon EMR. Semelhante ao Apache Hadoop, o Spark é um sistema de processamento distribuído de código-fonte aberto comumente utilizado para workloads de big data. No entanto, o Spark tem várias diferenças notáveis em relação ao MapReduce Hadoop. O Spark possui um mecanismo de execução otimizado para gráficos acíclicos dirigidos (DAG) e faz ativamente o cache dos dados na memória, o que pode aumentar o desempenho, especialmente para determinados algoritmos e consultas interativas.

O Spark oferece suporte nativo a aplicativos escritos em Scala, Python e Java. Ele também inclui várias bibliotecas totalmente integradas para SQL (Spark), aprendizado de máquina (MLlib), processamento de fluxo (streaming Spark) e processamento gráfico (GraphX). Essas ferramentas facilitam o aproveitamento da estrutura do Spark para uma grande variedade de casos de uso.

Você pode instalar o Spark em um cluster do Amazon EMR com outras aplicações do Hadoop e ele também pode aproveitar o sistema de arquivos do Amazon EMR (EMRFS) para acessar dados diretamente no Amazon S3. O Hive também é integrado ao Spark para que você possa usar um HiveContext objeto para executar scripts do Hive usando o Spark. Um contexto do Hive está incluído no spark-shell como `sqlContext`.

Para ver um exemplo de tutorial sobre como configurar um cluster do EMR com o Spark e analisar um conjunto de dados de amostra, consulte Tutorial: Getting started with Amazon EMR no blog de notícias. AWS

> **Important**
>
> A versão 2.3.1 do Apache Spark, disponível desde a versão 5.16.0 do Amazon EMR, aborda CVE-2018-8024 e CVE-2018-1334. Recomendamos que você migre as versões anteriores do Spark para a versão 2.3.1 ou posteriores.

A tabela a seguir lista a versão do Spark incluída na versão mais recente da série 7.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Spark.

Para a versão dos componentes instalados com o Spark nesta versão, consulte Versões de componentes da versão 7.9.0.

Informações sobre a versão do Spark para o emr-7.9.0

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-7.9.0	Spark 3.5.5	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

A tabela a seguir lista a versão do Spark incluída na versão mais recente da série 6.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Spark.

Para obter a versão dos componentes instalados com o Spark nessa versão, consulte Release 6.15.0 Component Versions.

Informações de versão do Spark para o emr-6.15.0

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-6.15.0	Spark 3.4.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
		kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

> **Note**
>
> A versão 6.8.0 do Amazon EMR vem com a versão 3.3.0 do Apache Spark. Esta versão do Spark usa o Apache Log4j 2 e o arquivo `log4j2.properties` para configurar o Log4j nos processos do Spark. Se você usar o Spark no cluster ou criar clusters do EMR com parâmetros de configuração personalizados e quiser atualizar para a versão 6.8.0 do Amazon EMR, deverá migrar para a nova classificação de configuração `spark-log4j2` e para o formato de chave do Apache Log4j 2. Para obter mais informações, consulte Migrar do Apache Log4j 1.x para Log4j 2.x.

A tabela a seguir lista a versão do Spark incluída na versão mais recente da série 5.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Spark.

Para obter a versão dos componentes instalados com o Spark nesse lançamento, consulte Release 5.36.2 Component Versions.

Informações de versão do Spark para o emr-5.36.2

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.36.2	Spark 2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode,

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
		hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Tópicos

- Criação de um cluster com o Apache Spark
- Execução de aplicações do Spark com o Docker no Amazon EMR 6.x
- Use o catálogo do AWS Glue Data Catalog com o Spark no Amazon EMR
- Trabalhando com uma hierarquia de vários catálogos no AWS Glue Data Catalog com o Spark no Amazon EMR
- Configurar o Spark
- Otimizar a performance do Spark
- Armazenamento em Cache de Fragmentos Resultantes do Spark
- Uso do RAPIDS Accelerator do NVIDIA para Apache Spark
- Acessar o shell do Spark
- Use o Amazon SageMaker Spark para aprendizado de máquina
- Escrever uma aplicação do Spark
- Melhorar a performance do Spark com Amazon S3
- Adicionar uma etapa do Spark
- Visualizar o histórico de aplicações do Spark
- Acesse a web do Spark UIs

- Uso do conector do Amazon Kinesis Data Streams para streaming estruturado do Spark
- Usar a integração do Amazon Redshift para Apache Spark com o Amazon EMR
- Histórico de versões do Spark

Criação de um cluster com o Apache Spark

O procedimento a seguir cria um cluster com o Spark instalado usando Opções rápidas no console do Amazon EMR.

Como alternativa, use as Opções avançadas para personalizar ainda mais sua configuração de cluster ou para enviar etapas para instalar aplicações de forma programática e depois executar aplicações personalizadas. Com qualquer uma das opções de criação de cluster, você pode optar por usar o AWS Glue como seu metastore do Spark SQL. Consulte Use o catálogo do AWS Glue Data Catalog com o Spark no Amazon EMR para obter mais informações.

Para iniciar um cluster com o Spark instalado

1. Abra o console do Amazon EMR em https://console.aws.amazon.com /emr.
2. Escolha Criar cluster para usar Opções rápidas.
3. Insira um Nome de cluster. O nome do cluster não pode conter os caracteres <, >, $, | ou ` (crase).
4. Em Configuração de software, escolha uma opção de Versão.
5. Em Aplicações, escolha o pacote de aplicações do Spark.
6. Selecione outras opções conforme necessário e escolha Create cluster (Criar cluster).

> **Note**
>
> Para configurar o Spark quando estiver criando o cluster, consulte Configurar o Spark.

Para iniciar um cluster com o Spark instalado usando o AWS CLI

- Crie o cluster usando o comando a seguir.

```
aws emr create-cluster --name "Spark cluster" --release-label emr-7.9.0 --
applications Name=Spark \
```

```
--ec2-attributes KeyName=myKey --instance-type m5.xlarge --instance-count 3 --use-
default-roles
```

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

Iniciar um cluster com o Spark instalado usando o SDK para Java

Especifique o Spark como um aplicativo com `SupportedProductConfig` usado em `RunJobFlowRequest`.

- O seguinte exemplo mostra como criar um cluster com o Spark usando Java.

```
import com.amazonaws.AmazonClientException;
import com.amazonaws.auth.AWSCredentials;
import com.amazonaws.auth.AWSStaticCredentialsProvider;
import com.amazonaws.auth.profile.ProfileCredentialsProvider;
import com.amazonaws.services.elasticmapreduce.AmazonElasticMapReduce;
import com.amazonaws.services.elasticmapreduce.AmazonElasticMapReduceClientBuilder;
import com.amazonaws.services.elasticmapreduce.model.*;
import com.amazonaws.services.elasticmapreduce.util.StepFactory;

public class Main {

        public static void main(String[] args) {
                AWSCredentials credentials_profile = null;
                try {
                        credentials_profile = new
 ProfileCredentialsProvider("default").getCredentials();
                } catch (Exception e) {
                        throw new AmazonClientException(
                                        "Cannot load credentials from .aws/
credentials file. " +
                                                        "Make sure that the
 credentials file exists and the profile name is specified within it.",
                                        e);
```

```
                }

                AmazonElasticMapReduce emr =
 AmazonElasticMapReduceClientBuilder.standard()
                                .withCredentials(new
 AWSStaticCredentialsProvider(credentials_profile))
                                .withRegion(Regions.US_WEST_1)
                                .build();

                // create a step to enable debugging in the AWS Management Console
                StepFactory stepFactory = new StepFactory();
                StepConfig enabledebugging = new StepConfig()
                                .withName("Enable debugging")
                                .withActionOnFailure("TERMINATE_JOB_FLOW")

 .withHadoopJarStep(stepFactory.newEnableDebuggingStep());

                Application spark = new Application().withName("Spark");

                RunJobFlowRequest request = new RunJobFlowRequest()
                                .withName("Spark Cluster")
                                .withReleaseLabel("emr-5.20.0")
                                .withSteps(enabledebugging)
                                .withApplications(spark)
                                .withLogUri("s3://path/to/my/logs/")
                                .withServiceRole("EMR_DefaultRole")
                                .withJobFlowRole("EMR_EC2_DefaultRole")
                                .withInstances(new JobFlowInstancesConfig()
                                                .withEc2SubnetId("subnet-12ab3c45")
                                                .withEc2KeyName("myEc2Key")
                                                .withInstanceCount(3)

 .withKeepJobFlowAliveWhenNoSteps(true)
                                                .withMasterInstanceType("m4.large")

 .withSlaveInstanceType("m4.large"));
                RunJobFlowResult result = emr.runJobFlow(request);
                System.out.println("The cluster ID is " + result.toString());
        }
}
```

Execução de aplicações do Spark com o Docker no Amazon EMR 6.x

Com o Amazon EMR 6.0.0, os aplicativos Spark podem usar contêineres Docker para definir suas dependências de biblioteca, em vez de instalar dependências nas instâncias individuais da Amazon no cluster. EC2 Para executar o Spark com o Docker, primeiro é necessário configurar o registro do Docker e definir parâmetros adicionais ao enviar um aplicativo do Spark. Para obter mais informações, consulte Configurar a integração do Docker.

Quando o aplicativo é enviado, o YARN invoca o Docker para extrair a imagem especificada do Docker e executar o aplicativo Spark dentro de um contêiner do Docker. Isso permite definir e isolar dependências facilmente. Isso reduz o tempo de bootstrapping ou de preparação de instâncias no cluster do Amazon EMR com as bibliotecas necessárias para a execução do trabalho.

Considerações ao executar o Spark com o Docker

Ao executar o Spark com o Docker, verifique se os seguintes pré-requisitos são atendidos:

- O pacote `docker` e a CLI são instalados somente nos nós core e de tarefa.
- Nas versões 6.1.0 r posteriores do Amazon EMR, você pode, como alternativa, instalar o Docker em um nó primário usando os comandos a seguir.
  - ```
    sudo yum install -y docker
    sudo systemctl start docker
    ```
- O comando `spark-submit` sempre deve ser executado em uma instância primária no cluster do Amazon EMR.
- Os registros do Docker usados para resolver imagens do Docker devem ser definidos usando a API de classificação com a chave de classificação `container-executor` para definir parâmetros adicionais ao executar o cluster:
  - `docker.trusted.registries`
  - `docker.privileged-containers.registries`
- Para executar um aplicativo do Spark em um contêiner do Docker, as seguintes opções de configuração são necessárias:
  - `YARN_CONTAINER_RUNTIME_TYPE=docker`
  - `YARN_CONTAINER_RUNTIME_DOCKER_IMAGE={DOCKER_IMAGE_NAME}`
    ```

- Ao usar o Amazon ECR para recuperar imagens do Docker, é necessário configurar o cluster para ele se autenticar. Para fazer isso, é necessário usar a seguinte opção de configuração:
  - YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG={DOCKER_CLIENT_CONFIG_PATH_ON
- Nas versões 6.1.0 e posteriores do Amazon EMR, você não precisa usar o comando listado `YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG={DOCKER_CLIENT_CONFIG_PATH_ON_HDFS` quando o recurso de autenticação automático do ECR está habilitado.
- Qualquer imagem do Docker usada com o Spark deve ter o Java instalado na imagem do Docker.

Para obter mais informações sobre os pré-requisitos, consulte Configurar a integração do Docker.

Criar uma imagem do Docker

As imagens do Docker são criadas usando um arquivo do Docker, que define os pacotes e a configuração a serem incluídos na imagem. Os dois exemplos a seguir são usados por Dockerfiles e PySpark SparkR.

PySpark Dockerfile

As imagens do Docker criadas a partir desse Dockerfile incluem o Python 3 e o pacote Python. NumPy Esse arquivo do Docker usa o Amazon Linux 2 e o Amazon Corretto JDK 8.

```
FROM amazoncorretto:8

RUN yum -y update
RUN yum -y install yum-utils
RUN yum -y groupinstall development

RUN yum list python3*
RUN yum -y install python3 python3-dev python3-pip python3-virtualenv

RUN python -V
RUN python3 -V

ENV PYSPARK_DRIVER_PYTHON python3
ENV PYSPARK_PYTHON python3

RUN pip3 install --upgrade pip
RUN pip3 install numpy pandas

RUN python3 -c "import numpy as np"
```

Arquivo do Docker do SparkR

As imagens do Docker criadas com esse arquivo do Docker incluem R e o pacote randomForest CRAN. Esse arquivo do Docker inclui o Amazon Linux 2 e o Amazon Corretto JDK 8.

```
FROM amazoncorretto:8

RUN java -version

RUN yum -y update
RUN amazon-linux-extras install R4

RUN yum -y install curl hostname

#setup R configs
RUN echo "r <- getOption('repos'); r['CRAN'] <- 'http://cran.us.r-project.org';
 options(repos = r);" > ~/.Rprofile

RUN Rscript -e "install.packages('randomForest')"
```

Para obter mais informações sobre a sintaxe do arquivo do Docker, consulte a documentação de referência do arquivo do Docker.

Usar imagens do Docker do Amazon ECR

O Amazon Elastic Container Registry (Amazon ECR) é um registro de contêiner do Docker totalmente gerenciado que facilita o armazenamento, o gerenciamento e a implantação de imagens de contêiner do Docker. Ao ser usado o Amazon ECR, o cluster deve ser configurado para confiar na instância do ECR e você deve configurar a autenticação para que o cluster use imagens do Docker do Amazon ECR. Para obter mais informações, consulte Configurar o YARN para acessar o Amazon ECR.

Para garantir que os hosts do Amazon EMR possam acessar as imagens armazenadas no Amazon ECR, o cluster deve ter as permissões da política `AmazonEC2ContainerRegistryReadOnly` associada ao perfil de instância. Para obter mais informações, consulte Política `AmazonEC2ContainerRegistryReadOnly`.

Nesse exemplo, o cluster deve ser criado com a configuração adicional a seguir para garantir que o registro do Amazon ECR seja confiável. Substitua o *`123456789123.dkr.ecr.us-east-1.amazonaws.com`* endpoint pelo seu endpoint Amazon ECR.

```
[
  {
    "Classification": "container-executor",
    "Configurations": [
      {
        "Classification": "docker",
        "Properties": {
          "docker.privileged-containers.registries":
 "local,centos,123456789123.dkr.ecr.us-east-1.amazonaws.com",
          "docker.trusted.registries": "local,centos,123456789123.dkr.ecr.us-
east-1.amazonaws.com"
        }
      }
    ],
    "Properties": {}
  }
]
```

Usando PySpark com o Amazon ECR

O exemplo a seguir usa o PySpark Dockerfile, que será marcado e carregado no Amazon ECR. Depois de fazer o upload do Dockerfile, você pode executar o PySpark trabalho e consultar a imagem do Docker do Amazon ECR.

Depois de iniciar o cluster, use o SSH para se conectar a um nó principal e execute os comandos a seguir para criar a imagem local do Docker a partir do exemplo do PySpark Dockerfile.

Primeiro, crie um diretório e um arquivo do Docker.

```
mkdir pyspark
vi pyspark/Dockerfile
```

Cole o conteúdo do PySpark Dockerfile e execute os comandos a seguir para criar uma imagem do Docker.

```
sudo docker build -t local/pyspark-example pyspark/
```

Crie o repositório `emr-docker-examples` do ECR para os exemplos.

```
aws ecr create-repository --repository-name emr-docker-examples
```

Marque e carregue a imagem criada localmente no ECR, substituindo-a por seu *123456789123.dkr.ecr.us-east-1.amazonaws.com* endpoint de ECR.

```
sudo docker tag local/pyspark-example 123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-
docker-examples:pyspark-example
sudo docker push 123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-
examples:pyspark-example
```

Use SSH para se conectar ao nó primário e prepare um script do Python com o nome de arquivo `main.py`. Cole o conteúdo a seguir no arquivo `main.py` e salve-o.

```
from pyspark.sql import SparkSession
spark = SparkSession.builder.appName("docker-numpy").getOrCreate()
sc = spark.sparkContext

import numpy as np
a = np.arange(15).reshape(3, 5)
print(a)
```

No Amazon EMR 6.0.0, para enviar o trabalho, consulte o nome da imagem do Docker. Defina os parâmetros de configuração adicionais para garantir que a execução do trabalho use o Docker como o runtime. Ao ser usado o Amazon ECR, o `YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG` deve fazer referência ao arquivo `config.json` contendo as credenciais usadas para fazer a autenticação no Amazon ECR.

```
DOCKER_IMAGE_NAME=123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-
examples:pyspark-example
DOCKER_CLIENT_CONFIG=hdfs:///user/hadoop/config.json
spark-submit --master yarn \
--deploy-mode cluster \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG=
$DOCKER_CLIENT_CONFIG \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG=
$DOCKER_CLIENT_CONFIG \
--num-executors 2 \
main.py -v
```

Nas versões 6.1.0 e superiores do Amazon EMR, para enviar o trabalho, consulte o nome da imagem do Docker. Quando a autenticação automática do ECR estiver habilitada, execute o comando a seguir.

```
DOCKER_IMAGE_NAME=123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-
examples:pyspark-example
spark-submit --master yarn \
--deploy-mode cluster \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--num-executors 2 \
main.py -v
```

Quando o trabalho for concluído, anote o ID do aplicativo YARN e use o comando a seguir para obter a saída do PySpark trabalho.

```
yarn logs --applicationId application_id | grep -C2 '\[\['
LogLength:55
LogContents:
[[ 0  1  2  3  4]
 [ 5  6  7  8  9]
 [10 11 12 13 14]]
```

Usar o SparkR com o Amazon ECR

O exemplo a seguir usa o arquivo do Docker do SparkR, que será marcado e carregado no ECR. Depois de carregar o Dockerfile, você pode executar o trabalho do SparkR e consultar a imagem do Docker no Amazon ECR.

Após executar o cluster, use o SSH para se conectar a um nó core e execute os comandos a seguir a fim de criar a imagem local do Docker com base no exemplo de arquivo do Docker do SparkR.

Primeiro, crie um diretório e o arquivo do Docker.

```
mkdir sparkr
vi sparkr/Dockerfile
```

Cole os conteúdos do arquivo do Docker do SparkR e execute os comandos a seguir para criar uma imagem do Docker.

```
sudo docker build -t local/sparkr-example sparkr/
```

Marque e faça o upload da imagem criada localmente para o Amazon ECR, *123456789123.dkr.ecr.us-east-1.amazonaws.com* substituindo-a por seu endpoint Amazon ECR.

```
sudo docker tag local/sparkr-example 123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-
docker-examples:sparkr-example
sudo docker push 123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-
examples:sparkr-example
```

Use SSH para se conectar ao nó primário e prepare um script R com o nome `sparkR.R`. Cole o conteúdo a seguir no arquivo `sparkR.R`.

```
library(SparkR)
sparkR.session(appName = "R with Spark example", sparkConfig =
 list(spark.some.config.option = "some-value"))

sqlContext <- sparkRSQL.init(spark.sparkContext)
library(randomForest)
# check release notes of randomForest
rfNews()

sparkR.session.stop()
```

No Amazon EMR 6.0.0, para enviar o trabalho, consulte o nome da imagem do Docker. Defina os parâmetros de configuração adicionais para garantir que a execução do trabalho use o Docker como o runtime. Ao ser usado o Amazon ECR, o `YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG` deve fazer referência ao arquivo `config.json` contendo as credenciais usadas para fazer a autenticação no ECR.

```
DOCKER_IMAGE_NAME=123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-
examples:sparkr-example
DOCKER_CLIENT_CONFIG=hdfs:///user/hadoop/config.json
spark-submit --master yarn \
--deploy-mode cluster \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG=
$DOCKER_CLIENT_CONFIG \
```

```
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG=
$DOCKER_CLIENT_CONFIG \
sparkR.R
```

Nas versões 6.1.0 e superiores do Amazon EMR, para enviar o trabalho, consulte o nome da imagem do Docker. Quando a autenticação automática do ECR estiver habilitada, execute o comando a seguir.

```
DOCKER_IMAGE_NAME=123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-
examples:sparkr-example
spark-submit --master yarn \
--deploy-mode cluster \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
sparkR.R
```

Quando o trabalho for concluído, anote o ID do aplicativo YARN e use o comando a seguir para obter a saída do trabalho do SparkR. Esse exemplo inclui testes para garantir que a biblioteca randomForest, a versão instalada e as notas de release estejam disponíveis.

```
yarn logs --applicationId application_id | grep -B4 -A10 "Type rfNews"
randomForest 4.6-14
Type rfNews() to see new features/changes/bug fixes.
Wishlist (formerly TODO):

* Implement the new scheme of handling classwt in classification.

* Use more compact storage of proximity matrix.

* Allow case weights by using the weights in sampling?

=============================================================================
Changes in 4.6-14:
```

Use o catálogo do AWS Glue Data Catalog com o Spark no Amazon EMR

Usando o Amazon EMR versão 5.8.0 ou posterior, você pode configurar o Spark para usar o AWS Glue Data Catalog como seu metastore Apache Hive. Recomendamos essa configuração quando você precisa de um metastore persistente do Hive ou de um metastore do Hive compartilhado por diferentes clusters, serviços, aplicativos ou contas. AWS

Usando o Amazon EMR versão 6.5.0 ou posterior, você pode configurar o Spark para usar o AWS Glue Data Catalog com o Apache Iceberg.

Usando o Amazon EMR versão 7.5.0 ou posterior, você pode configurar o Spark para usar o AWS Glue Data Catalog como seu catálogo Iceberg REST.

AWS O Glue é um serviço de extração, transformação e carregamento (ETL) totalmente gerenciado que torna simples e econômico categorizar seus dados, limpá-los, enriquecê-los e movê-los de forma confiável entre vários armazenamentos de dados. O AWS Glue Data Catalog fornece um repositório unificado de metadados em uma variedade de fontes e formatos de dados, integrando-se ao Amazon EMR, bem como ao Amazon RDS, Amazon Redshift, Redshift Spectrum, Athena e qualquer aplicativo compatível com o metastore Apache Hive. AWS Os rastreadores do Glue podem inferir automaticamente o esquema dos dados de origem no Amazon S3 e armazenar os metadados associados no catálogo de dados. Para obter mais informações sobre o catálogo de dados, consulte Preenchendo o catálogo de dados do AWS Glue no Guia do desenvolvedor do AWS Glue.

Taxas separadas se aplicam ao AWS Glue. Há uma taxa mensal para armazenar e acessar os metadados no Catálogo de Dados, uma taxa horária cobrada por minuto pelas tarefas do AWS Glue ETL e pelo tempo de execução do rastreador, e uma taxa horária cobrada por minuto para cada endpoint de desenvolvimento provisionado. O Catálogo de Dados permite o armazenamento de até um milhão de objetos gratuitamente. Se você armazenar mais de um milhão de objetos, será cobrado uma taxa de USD$1 por cada 100.000 objetos em mais de um milhão. Um objeto no Catálogo de Dados é uma tabela, uma partição ou um banco de dados. Para obter mais informações, consulte Definição de preço do Glue.

> **Important**
>
> Se você criou tabelas usando o Amazon Athena ou o Amazon Redshift Spectrum antes de 14 de agosto de 2017, bancos de dados e tabelas são armazenados em um catálogo gerenciado pelo Athena, que é separado do Glue Data Catalog. AWS Para integrar o

Amazon EMR a essas tabelas, você deve fazer o upgrade para o AWS Glue Data Catalog. Para obter mais informações, consulte Atualização para o catálogo de dados AWS Glue no Guia do usuário do Amazon Athena.

Especificando o AWS Glue Data Catalog como metastore do Apache Hive

Você pode especificar o AWS Glue Data Catalog como metastore usando a API AWS Management Console AWS CLI, ou Amazon EMR. Ao usar a CLI ou a API, você usa a classificação de configuração do Spark para especificar o Catálogo de Dados. Além disso, com o Amazon EMR 5.16.0 e versões posteriores, você pode usar a classificação de configuração para especificar um catálogo de dados em outro. Conta da AWS Ao usar o console, você pode especificar o Catálogo de Dados usando Opções avançadas ou Opções rápidas.

> **Note**
>
> A opção de usar o AWS Glue Data Catalog também está disponível com o Zeppelin porque o Zeppelin é instalado com componentes do Spark.

Console

Para especificar o AWS Glue Data Catalog como a metastore do Apache Hive com o novo console

1. Faça login no e abra AWS Management Console o console do Amazon EMR em https://console.aws.amazon.com /emr.
2. Em Amazon EMR, EC2 no painel de navegação esquerdo, escolha Clusters e, em seguida, escolha Create cluster.
3. Em Pacote de aplicações, escolha Spark ou Personalizar. Se você personalizar o cluster, selecione Zeppelin ou Spark como uma das suas aplicações.
4. Em Configurações do Catálogo de Dados do AWS Glue, marque a caixa de seleção Usar para metadados de tabelas do Spark.
5. Escolha qualquer outra opção que se aplique ao cluster.
6. Para iniciar o cluster, escolha Criar cluster.

AWS CLI

Para especificar o AWS Glue Data Catalog como o metastore do Apache Hive com o AWS CLI

Para obter mais informações sobre como especificar uma classificação de configuração usando a AWS CLI API do Amazon EMR, consulte. Configurar aplicações

- Especifique o valor para `hive.metastore.client.factory.class` usando a classificação `spark-hive-site`, conforme mostrado neste exemplo:

```
[
  {
    "Classification": "spark-hive-site",
    "Properties": {
      "hive.metastore.client.factory.class":
 "com.amazonaws.glue.catalog.metastore.AWSGlueDataCatalogHiveClientFactory"
    }
  }
]
```

Para especificar um catálogo de dados em uma AWS conta diferente, adicione a `hive.metastore.glue.catalogid` propriedade conforme mostrado no exemplo a seguir. Substitua *`acct-id`* pela conta da AWS do Catálogo de Dados.

```
[
  {
    "Classification": "spark-hive-site",
    "Properties": {
      "hive.metastore.client.factory.class":
 "com.amazonaws.glue.catalog.metastore.AWSGlueDataCatalogHiveClientFactory",
      "hive.metastore.glue.catalogid": "acct-id"
    }
  }
]
```

Especificando o AWS Glue Data Catalog como o catálogo Apache Iceberg

Você pode especificar o AWS Glue Data Catalog como a implementação do catálogo Apache Iceberg ou o endpoint do catálogo Apache Iceberg REST, usando a, a ou a API do Amazon EMR AWS Management Console, ou na configuração de tempo de execução AWS CLI da sessão do Spark. Ao usar a CLI ou a API, você usa a classificação de configuração do Spark para especificar o Catálogo de Dados. Para obter mais detalhes, consulte Especificando o AWS Glue Data Catalog como o catálogo Apache Iceberg.

permissões do IAM

O perfil da EC2 instância de um cluster deve ter permissões do IAM para ações do AWS Glue. Além disso, se você habilitar a criptografia para objetos do AWS Glue Data Catalog, a função também deverá ter permissão para criptografar, descriptografar e gerar o AWS KMS key usado para criptografia.

Permissões para ações do AWS Glue

Se você usar o perfil de EC2 instância padrão para o Amazon EMR, nenhuma ação será necessária. A política `AmazonElasticMapReduceforEC2Role` gerenciada anexada ao `EMR_EC2_DefaultRole` permite todas as ações necessárias do AWS Glue. No entanto, se você especificar um perfil de EC2 instância e permissões personalizados, deverá configurar as ações apropriadas do AWS Glue. Use a política gerenciada `AmazonElasticMapReduceforEC2Role` como ponto de partida. Para obter mais informações, consulte Função de serviço para EC2 instâncias de cluster (perfil de EC2 instância) no Guia de gerenciamento do Amazon EMR.

Permissões para criptografar e descriptografar o Glue Data Catalog AWS

O perfil de instância precisa de permissão para criptografar e descriptografar dados usando a chave. Você não precisa configurar essas permissões se as duas instruções a seguir se aplicarem:

- Você ativa a criptografia para objetos do AWS Glue Data Catalog usando chaves gerenciadas para o AWS Glue.
- Você usa um cluster que está no mesmo catálogo Conta da AWS de dados do AWS Glue.

Caso contrário, você deverá adicionar a seguinte declaração à política de permissões anexada ao seu perfil de EC2 instância.

```
[
```

```
    {
        "Version": "2012-10-17",
        "Statement": [
            {
                "Effect": "Allow",
                "Action": [
                    "kms:Decrypt",
                    "kms:Encrypt",
                    "kms:GenerateDataKey"
                ],
                "Resource": "arn:aws:kms:region:acct-
id:key/12345678-1234-1234-1234-123456789012"
            }
        ]
    }
]
```

Para obter mais informações sobre a criptografia do AWS Glue Data Catalog, consulte Criptografando seu catálogo de dados no AWS Glue Developer Guide.

Permissões baseadas em recursos

Se você usa o AWS Glue em conjunto com o Hive, o Spark ou o Presto no Amazon EMR AWS , o Glue oferece suporte a políticas baseadas em recursos para controlar o acesso aos recursos do catálogo de dados. Esses recursos incluem bancos de dados, tabelas, conexões e funções definidas pelo usuário. Para obter mais informações, consulte Políticas baseadas em recursos no AWS Glue no Guia do desenvolvedor do AWS Glue.

Ao usar políticas baseadas em recursos para limitar o acesso ao AWS Glue a partir do Amazon EMR, o principal que você especifica na política de permissões deve ser o ARN da função associado ao perfil da EC2 instância que é especificado quando um cluster é criado. Por exemplo, para uma política baseada em recursos anexada a um catálogo, você pode especificar o ARN da função para a função de serviço padrão para EC2 instâncias de cluster, *EMR_EC2_DefaultRole* como a`Principal`, usando o formato mostrado no exemplo a seguir:

```
arn:aws:iam::acct-id:role/EMR_EC2_DefaultRole
```

*acct-id*Pode ser diferente do ID da conta AWS Glue. Isso possibilita o acesso de clusters do EMR em outras contas. Você pode especificar várias entidades principais, cada uma de uma conta diferente.

Considerações ao usar o Catálogo de Dados do AWS Glue

Considere os seguintes itens ao usar o AWS Glue Data Catalog como metastore do Apache Hive com o Spark:

- Ter um banco de dados padrão sem um URI local causa falhas quando você cria uma tabela. Como alternativa, use a cláusula `LOCATION` para especificar o local do bucket, como `s3://amzn-s3-demo-bucket1`, ao usar `CREATE TABLE`. Ou crie tabelas em um banco de dados que não seja o banco de dados padrão.
- Não há suporte para renomear tabelas de dentro do AWS Glue.
- Quando você cria uma tabela do Hive sem especificar um `LOCATION`, os dados da tabela são armazenados no local especificado pela propriedade `hive.metastore.warehouse.dir`. Por padrão, esse é um local no HDFS. Se outro cluster precisar acessar a tabela, ele apresentará falha, a menos que tenha permissões adequadas para o cluster que criou a tabela. Além disso, como o armazenamento do HDFS é transitório, se o cluster for encerrado, os dados da tabela serão perdidos e a tabela precisará ser recriada. Recomendamos que você especifique a `LOCATION` no Amazon S3 ao criar uma tabela do Hive usando o Glue. AWS Como alternativa, você pode usar a classificação de configuração `hive-site` para especificar um local no Amazon S3 para `hive.metastore.warehouse.dir`, que se aplica a todas as tabelas do Hive. Se uma tabela for criada em um local do HDFS e o cluster que a criou ainda estiver em execução, você poderá atualizar o local da tabela para o Amazon S3 a partir AWS do Glue. Para obter mais informações, consulte Como trabalhar com tabelas no console do AWS Glue no Guia do desenvolvedor do AWS Glue.
- Valores de partição contendo aspas e apóstrofos não são compatíveis, por exemplo, `PARTITION (owner="Doe's").`
- As estatísticas de coluna são têm suporte para o emr-5.31.0 e versões posteriores.
- O uso da autorização do Hive não é compatível. Como alternativa, considere usar as Políticas baseadas em recursos do AWS Glue. Para obter mais informações, consulte Usar políticas baseadas em recursos para o Amazon EMR Access to AWS Glue Data Catalog.

Considere o seguinte ao usar o AWS Glue Data Catalog como catálogo REST do Apache Iceberg com o Spark:

- Se você usar o catálogo de sessões do Spark com o Iceberg, descrito emDiferenças de configuração quando você usa o Iceberg versus SparkCatalog SparkSessionCatalog, deverá

configurar o AWS Glue Data Catalog como a metastore do Apache Hive, além de configurar o AWS Glue Data Catalog como o catálogo REST do Apache Iceberg.

- O endpoint IRC do AWS Glue Data Catalog suporta somente o esquema de autenticação Amazon SigV4. OAuth não é suportado. Para OAuth usuários, use o IAM Identity Center para configurar o acesso. Consulte Conexão do Lake Formation com o IAM Identity Center.
- O catálogo REST do AWS Glue Iceberg não oferece suporte a todas as operações em código aberto.

Trabalhando com uma hierarquia de vários catálogos no AWS Glue Data Catalog com o Spark no Amazon EMR

Você pode registrar seu cluster do Amazon EMR para acessar o AWS Glue Data Catalog, que disponibiliza tabelas e outros recursos de catálogo para vários consumidores. AWS O Glue Data Catalog suporta uma hierarquia de vários catálogos, que unifica seus dados nos data lakes do Amazon S3. Ele também fornece uma API de metastore do Hive e uma API REST Apache Iceberg de código aberto para acessar os dados. Esses recursos estão disponíveis para o Amazon EMR e outros serviços, como Amazon Athena o Amazon Redshift.

Como os recursos do catálogo são organizados

Ao criar recursos no AWS Glue Data Catalog, você pode acessá-los de qualquer mecanismo SQL compatível com a API REST Apache Iceberg ou o metastore Hive. AWS Lake Formation gerencia as permissões.

No AWS Glue Data Catalog, os dados são organizados em uma hierarquia lógica de catálogos, bancos de dados e tabelas:

- Catálogo — Um contêiner lógico que contém objetos de um armazenamento de dados, como esquemas ou tabelas.
- Catálogo para armazenar tabelas do Redshift Managed Storage (RMS) — Ao gerenciar catálogos para armazenar tabelas RMS, você pode acessar essas tabelas usando o Iceberg.
- Banco de dados — organiza objetos de dados, como tabelas e visualizações, em um catálogo.
- Tabelas e visualizações — objetos de dados em um banco de dados que fornecem uma camada de abstração com um esquema compreensível. Eles fornecem uma camada para acessar os dados subjacentes, que podem estar em vários formatos e em diferentes locais.

Configurando um catálogo de dados para uso com o Amazon EMR

Para começar, você configura o catálogo para oferecer suporte às ferramentas do Amazon EMR. O AWS Glue Data Catalog oferece compatibilidade com o Hive metastore e com o Iceberg REST. APIs

Configurando o Amazon EMR com um metastore Hive

Para obter informações sobre como configurar isso, consulte Suporte do AWS Glue Data Catalog para tarefas do Spark no Guia do usuário do AWS Glue. Este tópico descreve como configurar o AWS Glue Data Catalog como um metastore do Hive e disponibilizá-lo como um endpoint. Além disso, há documentação do Amazon EMR disponível que mostra como especificar o AWS Glue Data Catalog como metastore do Spark, em Use the AWS Glue Data Catalog as the Apache Hive metastore for Spark.

Permissões para acessar recursos no AWS Glue Data Catalog

Esta seção descreve os requisitos da política do IAM para usar as ferramentas do Amazon EMR com dados do catálogo. Depois de registrar seu cluster no AWS Glue Data Catalog, você precisa das seguintes permissões para descobrir a criação e as alterações no catálogo de dados criado posteriormente:

- cola: GetCatalog
- cola: GetCatalogs
- conjuntos: AssumeRole
- conjuntos: TagSession
- conjuntos: SetContext
- conjuntos: SetSourceIdentity

Na maioria dos casos, ao atribuir permissões, recomendamos criar uma função do IAM e atribuir permissões a ela.

Além disso, para consultar dados do catálogo, você deve definir permissões para o catálogo de dados usando AWS Lake Formation. Para obter mais informações sobre como definir permissões para catálogos de dados em AWS Lake Formation, consulte Conceder e revogar permissões em recursos do Catálogo de Dados.

Depois de criar e configurar seu cluster e definir permissões em seus objetos de catálogo, você pode enviar trabalhos para consultar e processar dados.

Configure o Spark para acessar uma hierarquia de vários catálogos no AWS Glue Data Catalog

Com o EMR 7.5, você pode configurar o Spark para usar a hierarquia de vários catálogos do AWS Glue. Uma hierarquia de vários catálogos permite que você:

- Traga seus dados do Redshift Managed Storage (RMS), como tabelas, visualizações e visualizações materializadas de data warehouses existentes do Amazon Redshift, para o Glue Data Catalog. AWS Você pode consultar esses objetos usando o EMR on e o EMR EC2 Serverless.
- Crie catálogos RMS, AWS Glue Data Catalog e armazene dados no RMS usando o ZeroETL e consulte os dados com mecanismos de consulta compatíveis com o Iceberg.
- Crie tabelas Iceberg gerenciadas no AWS Glue Data Catalog com gerenciamento de armazenamento completo que inclui compactação, instantâneos e retenção.

Conectando-se a vários catálogos ao inicializar uma sessão do Spark

Os exemplos a seguir demonstram como usar o shell interativo do Spark, o envio do Spark ou os notebooks Amazon EMR para trabalhar AWS com a hierarquia de vários catálogos do Glue.

spark-shell

1. Conecte-se ao nó principal usando SSH. Para obter mais informações, consulte Conectar-se ao nó principal usando SSH no Guia de gerenciamento do Amazon EMR.
2. Digite o seguinte comando para iniciar o shell do Spark. Para usar a PySpark concha, `spark-shell` substitua por`pyspark`.

```
spark-shell \
    --conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
    --conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket/prefix/
    --conf spark.sql.catalog.my_catalog.type=glue \
    --conf spark.sql.catalog.my_catalog.glue.id=Glue RMS catalog ID \
    --conf spark.sql.defaultCatalog=my_catalog \
    --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

spark-submit

1. Conecte-se ao nó principal usando SSH. Para obter mais informações, consulte Conectar-se ao nó principal usando SSH no Guia de gerenciamento do Amazon EMR.
2. Digite o comando a seguir para iniciar a sessão do Spark para o Spark.

```
spark-submit \
--conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
--conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket1/prefix \
--conf spark.sql.catalog.my_catalog.type=glue \
--conf spark.sql.catalog.my_catalog.glue.id=Glue RMS catalog ID \
--conf spark.sql.defaultCatalog=my_catalog \
--conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

EMR Studio notebooks

Para inicializar uma sessão do Spark usando Cadernos do EMR Studio, configure a sessão do Spark usando o comando mágico `%%configure` no Caderno do Amazon EMR, como no exemplo a seguir. Para obter mais informações, consulte Use EMR Notebooks magics no Guia de gerenciamento do Amazon EMR.

```
%%configure -f{
"conf":{
    "spark.sql.catalog.my_catalog":"org.apache.iceberg.spark.SparkCatalog",
    "spark.sql.catalog.my_catalog.type":"glue",
    "spark.sql.catalog.my_catalog.glue.id":"Glue RMS catalog ID",
    "spark.sql.catalog.my_catalog.warehouse":"s3://amzn-s3-demo-bucket1/prefix/",
    "spark.sql.defaultCatalog", "my_catalog",

 "spark.sql.extensions":"org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions"
    }
}
```

CLI

Para inicializar uma sessão do Spark usando a CLI, execute o exemplo a seguir. Para obter mais informações sobre como especificar uma classificação de configuração usando a AWS CLI API do Amazon EMR, consulte Configurar aplicativos.

```
[
  {
    "Classification": "spark-defaults",
    "Properties": {
      "spark.sql.catalog.my_catalog":"org.apache.iceberg.spark.SparkCatalog",
      "spark.sql.catalog.my_catalog.type":"glue",
      "spark.sql.catalog.my_catalog.glue.id":"Glue RMS catalog ID",
      "spark.sql.catalog.my_catalog.warehouse":"s3://amzn-s3-demo-bucket1/prefix/",
      "spark.sql.defaultCatalog", "my_catalog",

 "spark.sql.extensions":"org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions"
    }
  }
]
```

Inicialize uma sessão do Spark no Redshift Managed Storage com o Glue Data Catalog AWS

O exemplo de comando a seguir inicializa uma sessão do Spark com o AWS Glue Data Catalog.

```
spark-sql \
  --conf spark.sql.catalog.rms=org.apache.iceberg.spark.SparkCatalog \
  --conf spark.sql.catalog.rms.type=glue \
  --conf spark.sql.catalog.rms.glue.id=Glue RMS catalog ID \
  --conf spark.sql.defaultCatalog=rms
  --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

O exemplo a seguir inicializa uma sessão do Spark usando a API REST Iceberg e o Redshift Managed Storage with Glue Data Catalog. AWS

```
spark-sql \
  --conf spark.sql.catalog.rms=org.apache.iceberg.spark.SparkCatalog \
  --conf spark.sql.catalog.rms.type=rest \
  --conf spark.sql.catalog.rms.warehouse=glue RMS catalog ID \
  --conf spark.sql.catalog.rms.uri=Glue endpoint URI/iceberg \
  --conf spark.sql.catalog.rms.rest.sigv4-enabled=true \
  --conf spark.sql.catalog.rms.rest.signing-name=glue \
  --conf spark.sql.defaultCatalog=rms \
  --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

Para obter mais informações sobre como usar uma hierarquia de vários catálogos do AWS Glue com o Spark Iceberg, consulte Usar um cluster Iceberg com o Spark.

Considerações e limitações para uma configuração de vários catálogos

- O uso de uma hierarquia de vários catálogos com o metastore Apache Hive não é suportado.
- O uso de uma hierarquia de vários catálogos com o Apache Iceberg não oferece suporte ao fallback para o metastore Apache Hive durante o uso. `SparkSessionCatalog`
- O EMR em EC2 clusters com a função Runtime não oferece suporte à hierarquia de vários catálogos.
- O EMR em EC2 clusters habilitados com AWS Lake Formation não oferece suporte à hierarquia de vários catálogos.

Configurar o Spark

Você pode configurar o Spark no Amazon EMR com classificações de configuração. Para obter mais informações sobre classificações de configuração, consulte Configurar aplicações.

As classificações de configuração para o Spark no Amazon EMR incluem o seguinte:

- **`spark`**: define a propriedade `maximizeResourceAllocation` como verdadeira ou falsa. Quando verdadeira, o Amazon EMR configura automaticamente as propriedades `spark-defaults` com base na configuração de hardware do cluster. Para obter mais informações, consulte Usar o `maximizeResourceAllocation`.
- **`spark-defaults`**: define valores no arquivo `spark-defaults.conf`. Para obter mais informações, consulte Configuração do Spark na documentação do Spark.
- **`spark-env`**: define valores no arquivo `spark-env.sh`. Para obter mais informações, consulte Variáveis de ambiente na documentação do Spark.
- **`spark-hive-site`**: define os valores no `hive-site.xml` para o Spark.
- **`spark-log4j`**— (Amazon EMR versões 6.7.x e inferiores) Define valores no arquivo. `log4j.properties` Para obter mais informações, consulte o arquivo log4j.properties.template no GitHub.
- **`spark-log4j2`**: (versões 6.8.0 e superiores do Amazon EMR) define valores no arquivo `log4j2.properties`. Para obter mais informações, consulte o arquivo log4j2.properties.template no GitHub.

- **`spark-metrics`**: define valores no arquivo `metrics.properties`. Para obter configurações e mais informações, consulte o arquivo metrics.properties.template no Github e as Métricas na documentação do Spark.

> **Note**
>
> Se você estiver migrando workloads do Spark para o Amazon EMR de outra plataforma, recomendamos que você teste suas workloads com o Padrões do Spark definidos pelo Amazon EMR antes de adicionar configurações personalizadas. A maioria dos clientes observa uma melhor performance com nossas configurações padrão.

Tópicos

- Padrões do Spark definidos pelo Amazon EMR
- Configurar a coleta de resíduos do Spark no Amazon EMR 6.1.0
- Usar o maximizeResourceAllocation
- Configurar o comportamento de desativação de nós
- Variável de ThriftServer ambiente Spark
- Alterar as configurações padrão do Spark
- Migrar do Apache Log4j 1.x para Log4j 2.x

Padrões do Spark definidos pelo Amazon EMR

A tabela a seguir mostra como o Amazon EMR define valores padrão no `spark-defaults` que afetam aplicações.

Padrões do Spark definidos pelo Amazon EMR

Configuração	Descrição	Valor padrão
`spark.executor.memory`	A quantidade de memória a ser usada por processo de executor. Por exemplo: 1g, 2g.	Essa configuração é determinada pelos tipos de instância core e de tarefa no cluster.

Configuração	Descrição	Valor padrão
`spark.executor.cores`	O número de núcleos para uso em cada executor.	Essa configuração é determinada pelos tipos de instância core e de tarefa no cluster.
`spark.dynamicAllocation.enabled`	Quando verdadeira, use alocação dinâmica de recursos para aumentar e diminuir a escala verticalmente do número de executores registrados em uma aplicação com base na workload.	`true` (com as versões 4.4.0 e superiores do Amazon EMR) **Note** O serviço de shuffle do Spark é automaticamente configurado pelo Amazon EMR.
`spark.sql.hive.advancedPartitionPredicatePushdown.enabled`	Quando verdadeiro, o pushdown avançado de predicados de partição para o metastore do Hive é habilitado.	`true`

Configuração	Descrição	Valor padrão
`spark.sql.hive.stringLikePartitionPredicatePushdown.enabled`	Envia os filtros `startsWith` , `contains` e `endsWith` para o metastore do Hive. **Note** O Glue não é compatível com o pushdown de predicados para `startsWith` , `contains` ou `endsWith`. Se você estiver usando o metastore do Glue e encontrar erros devido ao pushdown de predicados para essas funções, defina essa configuração como `false`.	`true`

Configurar a coleta de resíduos do Spark no Amazon EMR 6.1.0

Definir configurações personalizadas de coleta de resíduos com `spark.driver.extraJavaOptions` e `spark.executor.extraJavaOptions` resulta em falha na inicialização do driver ou do executor com o Amazon EMR 6.1 devido a uma configuração de coleta de resíduos conflitante com o Amazon EMR 6.1.0. Para o Amazon EMR 6.1.0, a configuração padrão de coleta de resíduos é definida por meio de `spark.driver.defaultJavaOptions` e `spark.executor.defaultJavaOptions`. Essa configuração só se aplica ao Amazon EMR 6.1.0. As opções da JVM não relacionadas à coleta de resíduos, como aquelas para configurar registro em log (`-verbose:class`), ainda podem ser configuradas por meio de `extraJavaOptions`. Para obter mais informações, consulte Propriedades das aplicações do Spark.

Usar o `maximizeResourceAllocation`

Para configurar os executores para usarem o máximo de recursos possível em cada nó em um cluster, defina `maximizeResourceAllocation` como `true` na classificação de configuração do `spark`. O `maximizeResourceAllocation` é específico para o Amazon EMR. Quando você habilita `maximizeResourceAllocation`, o Amazon EMR calcula os recursos máximos de computação e memória disponíveis para um executor em uma instância no grupo de instâncias centrais. Em seguida, ele define as configurações `spark-defaults` correspondentes com base nos valores máximos calculados.

O Amazon EMR calcula os recursos máximos de computação e memória disponíveis para um executor com base em um tipo de instância da frota de instâncias centrais. Como cada frota de instâncias pode ter diferentes tipos e tamanhos de instância em uma frota, a configuração do executor usada pelo Amazon EMR pode não ser a melhor para os clusters, por isso não recomendamos utilizar as configurações padrão com a alocação máxima de recursos. Defina configurações personalizadas para os clusters da frota de instâncias.

> **Note**
>
> Você não deve usar a `maximizeResourceAllocation` opção em clusters com outros aplicativos distribuídos, como HBase o. O Amazon EMR usa configurações personalizadas do YARN para aplicações distribuídas, que podem entrar em conflito com `maximizeResourceAllocation` e causar falhas nas aplicações do Spark.

Veja a seguir um exemplo de classificação de configuração do Spark com `maximizeResourceAllocation` definido como `true`.

```
[
  {
    "Classification": "spark",
    "Properties": {
      "maximizeResourceAllocation": "true"
    }
  }
]
```

Configurações configuradas em **`spark-defaults`** quando **`maximizeResourceAllocation`** está habilitado

Configuração	Descrição	Valor
spark.default.parallelism	O número padrão de partições é RDDs retornado por transformações como join reduceByKey, e paralelize quando não definido pelo usuário.	2X o número de cores de CPU disponíveis para contêineres do YARN.
spark.driver.memory	Quantidade de memória a ser usada para o processo do driver, ou seja, onde SparkContext é inicializado. (por exemplo, 1g, 2g).	A configuração é definida com base nos tipos de instância do cluster. No entanto, como a aplicação do driver do Spark pode ser executada na instância primária ou em uma das instâncias core (por exemplo, nos modos de cliente do YARN e de cluster, respectivamente), isso é definido com base no menor tipo de instância desses dois grupos de instâncias.
spark.executor.memory	Quantidade de memória para uso por processo de executor. (por exemplo, 1g, 2g)	A configuração é definida com base nos tipos de instância core e de tarefa do cluster.
spark.executor.cores	O número de núcleos para uso em cada executor.	A configuração é definida com base nos tipos de instância core e de tarefa do cluster.
spark.executor.instances	O número de executores.	A configuração é definida com base nos tipos de instância core e de tarefa do cluster. Defina a menos que

Configuração	Descrição	Valor
		`spark.dynamicAllocation.enabled` seja explicitamente definido como true ao mesmo tempo.

Configurar o comportamento de desativação de nós

Com as versões 5.9.0 e superiores do Amazon EMR, o Spark no Amazon EMR inclui um conjunto de recursos para ajudar a garantir que o Spark lide tranquilamente com o encerramento de nós devido a um redimensionamento manual ou a uma solicitação de política de ajuste de escala automática. O Amazon EMR implementa um mecanismo de lista de negação no Spark criado com base no mecanismo de desativação do YARN. Esse mecanismo ajuda a garantir que não haja novas tarefas programadas em um nó que esteja sendo desativado, permitindo ao mesmo tempo que tarefas que já estão em execução sejam concluídas. Além disso, há recursos para ajudar a recuperar trabalhos do Spark com mais rapidez se os shuffle blocks forem perdidos no encerramento de um nó. O processo de recálculo é acionado mais cedo e otimizado para recalcular com mais rapidez e menos tentativas de fase, e é possível evitar falhas nos trabalhos provenientes de falhas de busca que são causadas por shuffle blocks ausentes.

> **Important**
>
> A configuração `spark.decommissioning.timeout.threshold` foi adicionada ao Amazon EMR versão 5.11.0 para melhorar a resiliência do Spark ao serem usadas instâncias spot. Nas versões anteriores, quando um nó usa uma instância spot e a instância é encerrada devido ao preço da oferta, o Spark pode não ser capaz de lidar com o encerramento com tranquilidade. Os trabalhos podem falhar e os novos cálculos de shuffle podem levar um tempo significativo. Por esse motivo, recomendamos o uso da versão 5.11.0, ou posterior, se você usar instâncias spot.

Configurações de desativação de nós do Spark

Configuração	Descrição	Valor padrão
`spark.blacklist.decommissioning.enabled`	Quando definido como `true`, o Spark coloca na lista de negação os nós que estão no estado `decommissioning` no YARN. O Spark não programa novas tarefas em executores que estejam em execução nesse nó. É permitido que as tarefas já em execução sejam concluídas.	`true`
`spark.blacklist.decommissioning.timeout`	O período em que um nó no estado `decommissioning` permanece na lista de negação. Por padrão, esse valor é definido como uma hora, que também é o padrão para `yarn.resourcemanager.decommissioning.timeout`. Para garantir que um nó permaneça na lista de negação por todo o período de desativação, defina esse valor como igual ou maior do que `yarn.resourcemanager.decommissioning.timeout`. Depois que o tempo limite de descomissionamento expirar, o nó passa para um estado `decommissioned` e o Amazon EMR pode encerrar a instância do nó. EC2 Se alguma tarefa	1h

Configuração	Descrição	Valor padrão
	ainda estiver em execução após o tempo limite expirar, ela será perdida ou eliminada e reprogramada em executores executados em outros nós.	

Configuração	Descrição	Valor padrão
`spark.decommissioning.timeout.threshold`	Disponível nas versões 5.11.0 ou posteriores do Amazon EMR. Especificado em segundos. Quando um nó muda para o estado de desativação, se o host é desativado em um período igual ou menor que esse valor, o Amazon EMR não apenas insere o nó em uma lista de negação, mas também limpa o estado do host (conforme especificado por `spark.resourceManager.cleanupExpiredHost`) sem esperar que o nó mude para um estado de desativação. Isso permite que o Spark lide melhor com o encerramento de instâncias spot, pois as instâncias spot são desativadas dentro de um tempo de espera de 20 segundos, independentemente do valor do `yarn.resourcemager.decommissioning.timeout` , o que pode não dar aos outros nós tempo suficiente para que eles leiam arquivos embaralhados.	`20s`

Configuração	Descrição	Valor padrão
`spark.resourceManager.cleanupExpiredHost`	Quando definido como `true`, o Spark cancela o registro de todos os dados em cache e os shuffle blocks armazenados nos executores nos nós que estejam no estado `decommissioned` . Isso acelera o processo de recuperação.	`true`
`spark.stage.attempt.ignoreOnDecommissionFetchFailure`	Quando definido como `true`, ajuda a evitar falhas do Spark nas fases e a falha no trabalho devido ao excesso de falhas de busca dos nós desativados. As buscas com falha de shuffle blocks de nó no estado `decommissioned` não são contabilizadas para o número máximo de falhas de busca consecutivas.	`true`

Variável de ThriftServer ambiente Spark

O Spark define a variável de ambiente da Porta do servidor Thrift do Hive, `HIVE_SERVER2_THRIFT_PORT`, como 10001.

Alterar as configurações padrão do Spark

Você altera os padrões em `spark-defaults.conf` usando a classificação de configuração `spark-defaults` ou a configuração `maximizeResourceAllocation` na classificação de configuração `spark`.

Os procedimentos a seguir mostram como modificar as configurações usando a CLI ou o console.

Criar um cluster com spark.executor.memory definido como 2g usando a CLI

- Crie um cluster com o Spark instalado e com `spark.executor.memory` definido como 2g usando o comando a seguir, que faz referência a um arquivo, `myConfig.json`, armazenado no Amazon S3.

```
aws emr create-cluster --release-label emr-7.9.0 --applications Name=Spark \
--instance-type m5.xlarge --instance-count 2 --service-role EMR_DefaultRole_V2
 --ec2-attributes InstanceProfile=EMR_EC2_DefaultRole --configurations https://
s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myConfig.json
```

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

`myConfig.json`:

```
[
    {
      "Classification": "spark-defaults",
      "Properties": {
        "spark.executor.memory": "2G"
      }
    }
  ]
```

Criar um cluster com spark.executor.memory definido como 2g usando o console

1. Navegue até o novo console do Amazon EMR e selecione Alternar para o console antigo na navegação lateral. Para obter mais informações sobre o que esperar ao alternar para o console antigo, consulte Usar o console antigo.
2. Escolha Create cluster (Criar cluster), Go to advanced options (Ir para opções avançadas).
3. Escolha Spark.
4. Em Edit software settings (Editar configurações de software), deixe Enter configuration (Inserir configuração) selecionado e insira a seguinte configuração:

```
classification=spark-defaults,properties=[spark.executor.memory=2G]
```

5. Selecione outras opções, escolha e Criar cluster.

Para definir maximizeResourceAllocation

- Crie um cluster com o Spark instalado e `maximizeResourceAllocation` definido como verdadeiro usando o AWS CLI, referenciando um arquivo`myConfig.json`, armazenado no Amazon S3.

```
aws emr create-cluster --release-label emr-7.9.0 --applications Name=Spark \
--instance-type m5.xlarge --instance-count 2 --service-role EMR_DefaultRole_V2
 --ec2-attributes InstanceProfile=EMR_EC2_DefaultRole --configurations https://
s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myConfig.json
```

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

`myConfig.json`:

```
[
  {
    "Classification": "spark",
    "Properties": {
      "maximizeResourceAllocation": "true"
    }
  }
]
```

> **Note**
>
> Com as versões 5.21.0 e posteriores do Amazon EMR, você pode substituir as configurações de cluster e especificar classificações de configuração adicionais para cada grupo de

instâncias em um cluster em execução. Você faz isso usando o console do Amazon EMR, o AWS Command Line Interface (AWS CLI) ou o AWS SDK. Para obter mais informações, consulte Supplying a Configuration for an Instance Group in a Running Cluster.

Migrar do Apache Log4j 1.x para Log4j 2.x

As versões 3.2.x e anteriores do Apache Spark usam o Apache Log4j 1.x herdado e o arquivo `log4j.properties` para configurar o Log4j nos processos do Spark. As versões 3.3.0 e posteriores do Apache Spark usam o Apache Log4j 2.x e o arquivo `log4j2.properties` para configurar o Log4j nos processos do Spark.

Se você tiver configurado o Apache Spark Log4j usando uma versão do Amazon EMR inferior à 6.8.0, deverá remover a classificação de configuração `spark-log4j` herdada e migrar para a classificação de configuração `spark-log4j2` e o formato da chave antes de poder atualizar para o Amazon EMR 6.8.0 ou posterior. A classificação `spark-log4j` herdada faz com que a criação do cluster apresente falha com um erro de `ValidationException` nas versões 6.8.0 e posteriores do Amazon EMR. Você não será cobrado por uma falha relacionada à incompatibilidade do Log4j, mas deverá remover a classificação de configuração `spark-log4j` extinta para continuar.

Para obter mais informações sobre a migração do Apache Log4j 1.x para o Log4j 2.x, consulte o Guia de migração do Apache Log4j e o Modelo do Spark Log4j 2 no GitHub.

> **Note**
>
> Com o Amazon EMR, o Apache Spark usa um arquivo `log4j2.properties` em vez do arquivo .xml descrito no Guia de migração do Apache Log4j. Além disso, não recomendamos o uso do método de ponte do Log4j 1.x para conversão para Log4j 2.x.

Otimizar a performance do Spark

O Amazon EMR fornece vários atributo de otimização de performance para o Spark. Este tópico explica detalhadamente cada recurso de otimização.

Para obter mais informações sobre como definir a configuração do Spark, consulte Configurar o Spark.

Execução adaptável de consultas

A execução adaptável de consultas é uma estrutura para a reotimização de planos de consulta com base nas estatísticas de runtime. Desde o Amazon EMR 5.30.0, as otimizações da execução adaptável de consultas do Apache Spark 3 descritas a seguir estão disponíveis no Runtime do Amazon EMR para Apache Spark 2.

- Conversão adaptável de junções
- Aglutinação adaptável de partições aleatórias

Conversão adaptável de junções

A conversão adaptativa de junções melhora o desempenho da consulta ao converter sort-merge-join operações em broadcast-hash-joins operações com base nos tamanhos de tempo de execução dos estágios da consulta. Broadcast-hash-joinstendem a ter um melhor desempenho quando um lado da junção é pequeno o suficiente para transmitir com eficiência sua saída para todos os executores, evitando assim a necessidade de trocar aleatoriamente e classificar os dois lados da junção. A conversão adaptativa de junções amplia a variedade de casos em que o Spark funciona automaticamente. broadcast-hash-joins

Esse recurso está habilitado por padrão. Ele pode ser desabilitado pela definição de `spark.sql.adaptive.enabled` como `false`, o que também desabilita a estrutura adaptável de execução de consultas. O Spark decide converter a em a broadcast-hash-join quando sort-merge-join a estatística do tamanho do tempo de execução de um dos lados da junção não excede`spark.sql.autoBroadcastJoinThreshold`, cujo padrão é 10.485.760 bytes (10 MiB).

Aglutinação adaptável de partições aleatórias

A aglutinação adaptável de partições aleatórias melhora a performance das consultas ao aglutinar pequenas partições aleatórias contíguas para evitar a sobrecarga de ter muitas tarefas pequenas. Isso permite que você configure antecipadamente um número maior de partições aleatórias iniciais que, em seguida, são reduzidas em runtime para um tamanho desejado, aumentando as probabilidades da existência de partições aleatórias distribuídas mais uniformemente.

Esse atributo é habilitado por padrão, a menos que `spark.sql.shuffle.partitions` seja definido explicitamente. Ele pode ser habilitado pela definição de `spark.sql.adaptive.coalescePartitions.enabled` como `true`. O número inicial de partições aleatórias e o tamanho da partição de destino podem ser ajustados com o uso das propriedades `spark.sql.adaptive.coalescePartitions.minPartitionNum` e

`spark.sql.adaptive.advisoryPartitionSizeInBytes`, respectivamente. Consulte a tabela a seguir para obter mais detalhes sobre as propriedades do Spark relacionadas a esse atributo.

Propriedades da partição de aglutinação adaptável do Spark

Propriedade	Valor padrão	Descrição
`spark.sql.adaptive.coalescePartitions.enabled`	verdadeiro, a menos que `spark.sql.shuffle.partitions` seja explicitamente definido	Quando verdadeiro e spark.sql.adaptive.enabled é verdadeiro, o Spark aglutina partições aleatórias contíguas de acordo com o tamanho de destino (especificado por `spark.sql.adaptive.advisoryPartitionSizeInBytes`), para evitar muitas tarefas pequenas.
`spark.sql.adaptive.advisoryPartitionSizeInBytes`	64 MB	O tamanho do comunicado, em bytes, da partição aleatória durante a aglutinação. Essa configuração só tem efeito quando `spark.sql.adaptive.enabled` e `spark.sql.adaptive.coalescePartitions.enabled` são ambos `true`.
`spark.sql.adaptive.coalescePartitions.minPartitionNum`	25	O número mínimo de partições aleatórias após a aglutinação. Essa configuração só tem efeito quando `spark.sql.adaptive.enabled` e `spark.sql.adaptive.coalescePartitions.enabled` são ambos `true`.

Propriedade	Valor padrão	Descrição
`spark.sql.adaptive.coalescePartitions.initialPartitionNum`	1000	O número inicial de partições aleatórias antes da aglutinação. Essa configuração só tem efeito quando `spark.sql.adaptive.enabled` e `spark.sql.adaptive.coalescePartitions.enabled` são ambos `true`.

Corte de partição dinâmico

O corte de partição dinâmico melhora o desempenho da tarefa selecionando as partições específicas com mais precisão em uma tabela que precisa ser lida e processada em uma consulta específica. Reduzindo a quantidade de dados que são lidos e processados, um tempo significativo é economizado na execução da tarefa. Com o Amazon EMR 5.26.0, esse atributo é habilitado por padrão. Com o Amazon EMR 5.24.0 e 5.25.0, é possível habilitar esse atributo ao definir a propriedade `spark.sql.dynamicPartitionPruning.enabled` do Spark no próprio Spark ou ao criar clusters.

Propriedades do corte de partição dinâmico do Spark

Propriedade	Valor padrão	Descrição
`spark.sql.dynamicPartitionPruning.enabled`	`true`	Quando verdadeiro, habilite o corte de partição dinâmico.
`spark.sql.optimizer.dynamicPartitionPruning.enforceBroadcastReuse`	`true`	Quando `true`, o Spark conduz uma verificação defensiva antes da execução da consulta para garantir que a reutilização de trocas de transmissão em filtros de corte dinâmico não seja interromp

Propriedade	Valor padrão	Descrição
		ida por regras de preparaçã o posteriores, como regras colunares definidas pelo usuário. Quando a reutiliza ção é interrompida e essa configuração é `true`, o Spark remove os filtros de corte dinâmico afetados para evitar problemas de performance e correção. Podem surgir problemas de correção quando a troca de transmiss ão do filtro de corte dinâmico produz resultados diferente s e inconsistentes da troca de transmissão da operação de junção correspondente. A definição dessa configura ção como `false` deve ser feita com cuidado; ela permite contornar cenários, como quando a reutilização é interrompida por regras colunares definidas pelo usuário. Quando a execução adaptável de consultas está habilitada, a reutilização da transmissão sempre é imposta.

Essa otimização melhora as capacidades existentes do Spark 2.4.2, que tem suporte apenas à redução de predicados estáticos que podem ser resolvidos na hora do planejamento.

Veja a seguir exemplos de redução de predicados estáticos no Spark 2.4.2.

```
partition_col = 5

partition_col IN (1,3,5)

partition_col between 1 and 3

partition_col = 1 + 3
```

O corte de partição dinâmico permite que o mecanismo do Spark deduza dinamicamente no runtime quais partições precisam ser lidas e quais podem ser eliminadas com segurança. Por exemplo, a consulta a seguir envolve duas tabelas: a tabela `store_sales` contém as vendas totais de todas as lojas e é particionada por região, e a tabela `store_regions` contém um mapeamento de regiões por país. As tabelas contêm dados sobre as lojas que estão distribuídas pelo mundo, mas estão consultado dados somente para a América do Norte.

```
select ss.quarter, ss.region, ss.store, ss.total_sales
from store_sales ss, store_regions sr
where ss.region = sr.region and sr.country = 'North America'
```

Sem o corte de partição dinâmico, essa consulta lerá todas as regiões antes de filtrar o subconjunto de regiões que corresponde aos resultados da subconsulta. Com o corte de partição dinâmico, essa consulta lerá e processará apenas as partições para as regiões retornadas na subconsulta. Isso economiza tempo e recursos ao ler menos dados do armazenamento e processar menos registros.

Nivelamento de subconsultas escalares

Esta otimização melhora o desempenho de consultas que têm subconsultas escalares na mesma tabela. Com o Amazon EMR 5.26.0, esse atributo é habilitado por padrão. Com o Amazon EMR 5.24.0 e 5.25.0, é possível habilitá-lo definindo a propriedade `spark.sql.optimizer.flattenScalarSubqueriesWithAggregates.enabled` do Spark no próprio Spark ou ao criar clusters. Quando essa propriedade é definida como true, o otimizador de consultas nivela as subconsultas escalares que usam a mesma relação, quando possível. As subconsultas escalares são niveladas enviando qualquer predicado presente na subconsulta para as funções agregadas e executando uma agregação, com todas as funções agregadas, por relação.

Veja a seguir um exemplo de consulta que se beneficia dessa otimização.

```
select (select avg(age) from students                    /* Subquery 1 */
```

```
                where age between 5 and 10) as group1,
       (select avg(age) from students                    /* Subquery 2 */
                where age between 10 and 15) as group2,
       (select avg(age) from students                    /* Subquery 3 */
                where age between 15 and 20) as group3
```

A otimização reescreve a consulta anterior como:

```
select c1 as group1, c2 as group2, c3 as group3
from (select avg (if(age between 5 and 10, age, null)) as c1,
             avg (if(age between 10 and 15, age, null)) as c2,
             avg (if(age between 15 and 20, age, null)) as c3 from students);
```

Observe que a consulta regravada lê a tabela do estudante somente uma vez, e os predicados das três subconsultas são enviados para a função `avg`.

DISTINCT antes de INTERSECT

Esta otimização otimiza junções usando INTERSECT. Com o Amazon EMR 5.26.0, esse atributo é habilitado por padrão. Com o Amazon EMR 5.24.0 e 5.25.0, é possível habilitá-lo definindo a propriedade `spark.sql.optimizer.distinctBeforeIntersect.enabled` do Spark no próprio Spark ou ao criar clusters. As consultas que usam INTERSECT são automaticamente convertidas para usar uma junção Left-Semi Join. Quando essa propriedade é definida como verdadeira, o otimizador de consultas envia o operador DISTINCT para os filhos de INTERSECT se detectar que o operador DISTINCT pode fazer com que a junção semi esquerda seja a em vez de a. BroadcastHashJoin SortMergeJoin

Veja a seguir um exemplo de consulta que se beneficia dessa otimização.

```
(select item.brand brand from store_sales, item
     where store_sales.item_id = item.item_id)
intersect
(select item.brand cs_brand from catalog_sales, item
     where catalog_sales.item_id = item.item_id)
```

Sem habilitar essa propriedade `spark.sql.optimizer.distinctBeforeIntersect.enabled`, a consulta será regravada do modo a seguir.

```
select distinct brand from
```

```
  (select item.brand brand from store_sales, item
     where store_sales.item_id = item.item_id)
left semi join
   (select item.brand cs_brand from catalog_sales, item
     where catalog_sales.item_id = item.item_id)
 on brand <=> cs_brand
```

Quando você habilita essa propriedade `spark.sql.optimizer.distinctBeforeIntersect.enabled`, a consulta pode ser regravada do modo a seguir.

```
select brand from
  (select distinct item.brand brand from store_sales, item
     where store_sales.item_id = item.item_id)
left semi join
   (select distinct item.brand cs_brand from catalog_sales, item
     where catalog_sales.item_id = item.item_id)
 on brand <=> cs_brand
```

Junção com filtro Bloom

Essa otimização pode melhorar o desempenho de algumas junções ao fazer uma pré-filtragem de um lado de uma junção usando um filtro Bloom gerado pelos valores do outro lado da junção. Com o Amazon EMR 5.26.0, esse atributo é habilitado por padrão. Com o Amazon EMR 5.25.0, é possível habilitar esse atributo definindo a propriedade `spark.sql.bloomFilterJoin.enabled` do Spark como `true` no próprio Spark ou ao criar clusters.

Veja a seguir um exemplo de consulta que podem se beneficiar de um filtro Bloom.

```
select count(*)
from sales, item
where sales.item_id = item.id
and item.category in (1, 10, 16)
```

Quando esse recurso está habilitado, o filtro Bloom é criado de todos os IDs de itens cuja categoria está no conjunto de categorias que está sendo consultado. Ao verificar a tabela de vendas, o filtro Bloom é usado para determinar quais são as vendas de itens que definitivamente não estão no conjunto definido pelo filtro Bloom. Assim, o filtro pode excluir essas vendas identificadas o mais cedo possível.

Reordenação de junção otimizada

Essa otimização pode melhorar o desempenho de consultas reordenando junções que envolvem tabelas com filtros. Com o Amazon EMR 5.26.0, esse atributo é habilitado por padrão. Com o Amazon EMR 5.25.0, é possível habilitar esse atributo definindo o parâmetro de configuração `spark.sql.optimizer.sizeBasedJoinReorder.enabled` do Spark como verdadeiro. O comportamento padrão no Spark é unir tabelas da esquerda para a direita, conforme listado na consulta. Essa estratégia pode perder oportunidades de executar junções menores com filtros primeiro a fim de beneficiar junções mais caras posteriormente.

O exemplo de consulta a seguir relata todos os itens devolvidos de todas as lojas de um país. Sem a reordenação de junção otimizada, o Spark une as duas tabelas grandes `store_sales` e `store_returns` primeiro, depois, as une a `store` e, por fim, a `item`.

```
select ss.item_value, sr.return_date, s.name, i.desc,
from store_sales ss, store_returns sr, store s, item i
where ss.id = sr.id and ss.store_id = s.id and ss.item_id = i.id
and s.country = 'USA'
```

Com reordenação de junção otimizada o Spark une `store_sales` a `store` primeiro, pois `store` tem um filtro e é menor que `store_returns` e `broadcastable`. Em seguida, o Spark une a `store_returns` e, por fim, a `item`. Se `item` tivesse um filtro e fosse possível de ser transmitido, ele também se qualificaria para a reordenação, resultando em `store_sales` se unindo a `store`, depois a `item` e, por fim, a `store_returns`.

Armazenamento em Cache de Fragmentos Resultantes do Spark

As versões 6.6.0 e superiores do Amazon EMR incluem o recurso opcional Armazenamento em Cache de Fragmentos Resultantes, que automaticamente armazena em cache os fragmentos resultantes. Esses fragmentos resultantes são partes dos resultados de subárvores de consultas que são armazenadas em um bucket do Amazon S3 de sua escolha. Os fragmentos resultantes da consulta armazenados são reutilizados em execuções de consultas subsequentes, resultando em consultas mais rápidas.

O Result Fragment Caching analisa suas consultas do Spark e armazena em cache os fragmentos de resultados elegíveis no local especificado do S3. Em execuções subsequentes da consulta, os fragmentos resultantes utilizáveis da consulta são automaticamente detectados e obtidos no S3. O Armazenamento em Cache de Fragmentos Resultantes difere do Armazenamento em Cache de

Conjuntos Resultantes, em que as consultas subsequentes precisam corresponder exatamente à consulta original para retornar os resultados do cache. Quando usado para consultas repetidamente direcionadas a um subconjunto estático dos dados, o Armazenamento em Cache de Fragmentos Resultantes acelera significativamente a performance.

Considere a consulta a seguir, que conta pedidos até o ano de 2022:

```
select
    l_returnflag,
    l_linestatus,
    count(*) as count_order
from
    lineitem
where
    l_shipdate <= current_date
    and year(l_shipdate) == '2022'
group by
    l_returnflag,
    l_linestatus
```

Com o passar do tempo, essa consulta precisa ser executada todos os dias para informar o total de vendas do ano. Sem o Armazenamento em Cache de Fragmentos Resultantes, os resultados de todos os dias do ano precisam ser recalculados todos os dias. A consulta ficará mais lenta com o tempo e será mais lenta no final do ano, quando todos os 365 dias de resultados precisarão ser recalculados.

Ao ativar o Armazenamento em Cache de Fragmentos Resultantes, você usa os resultados de todos os dias anteriores do ano com base no cache. A cada dia, o atributo deve recalcular apenas um dia de resultados. Depois que o atributo computa o fragmento resultante, o atributo armazena o fragmento em cache. Como resultado, os tempos de consultas habilitadas para cache são curtos e permanecem constantes para cada consulta subsequente.

Habilitar o Armazenamento em Cache de Fragmentos Resultantes do Spark

Para habilitar o Armazenamento em Cache de Fragmentos Resultantes do Spark, execute as seguintes etapas:

1. Crie um bucket de cache no Amazon S3 e autorize o read/write acesso ao EMRFS. Para obter mais informações, consulte Autorizar acesso aos dados do EMRFS no Amazon S3.

2. Defina a configuração do Spark no Amazon EMR para habilitar o recurso.

```
spark.subResultCache.enabled = true
spark.subResultCache.fs.root.path = s3://&example-s3-bucket;/cache_dir/
```

3. Habilite o gerenciamento do ciclo de vida do S3 para que o bucket limpe automaticamente os arquivos de cache.
4. Opcionalmente, configure as maxBufferSize propriedades reductionRationThreshold e para ajustar ainda mais o recurso.

```
spark.sql.subResultCache.reductionRatioThreshold
spark.sql.subResultCache.maxBufferSize
```

Considerações ao usar o Armazenamento em Cache de Fragmentos Resultantes

A economia de custos obtida quando você usa resultados já armazenados em cache no Amazon S3 em vez de recalculá-los aumenta com o número de vezes que os mesmos resultados armazenados em cache podem ser usados. Consultas com verificações de grandes tabelas seguidas por filtros ou agregações de hash que reduzem o tamanho do resultado em um fator de pelo menos 8 (ou seja, uma proporção de pelo menos 8:1 em tamanho de entrada:resultados) serão as que mais se beneficiarão desse atributo. Quanto maior a taxa de redução entre a entrada e os resultados, maior a relação custo-benefício. Consultas com taxas de redução menores, mas que contêm etapas computacionais caras entre a verificação da tabela e o filtro ou as agregações, também serão beneficiadas, desde que o custo de produzir os resultados seja maior do que o custo de buscá-los no Amazon S3. Por padrão, o Armazenamento em Cache de Fragmentos Resultantes só entra em vigor quando detecta que uma taxa de redução será de pelo menos 8:1.

Quando suas consultas reutilizam repetidamente os resultados armazenados em cache, os benefícios desse atributo são maiores. Consultas em janelas contínuas e incrementais são bons exemplos. Por exemplo, uma consulta de janela contínua de 30 dias que já foi executada por 29 dias precisaria extrair apenas 1/30 dos dados de destino de sua fonte de entrada original e usaria fragmentos de resultados armazenados em cache nos 29 dias anteriores. Uma consulta de janela incremental se beneficiaria ainda mais, já que o início da janela permanece fixo: em cada invocação da consulta, uma porcentagem menor do processamento exigirá a leitura da fonte de entrada.

Veja a seguir considerações adicionais sobre o uso do Armazenamento em Cache de Fragmentos Resultantes:

- As consultas que não têm como alvo os mesmos dados com os mesmos fragmentos de consulta terão uma baixa taxa de acertos de cache e, portanto, não se beneficiarão desse atributo.
- Consultas com taxas de redução baixas que não contêm etapas de computação caras resultarão em resultados armazenados em cache que são aproximadamente tão caros de ler quanto de processar inicialmente.
- A primeira consulta sempre demonstrará uma pequena regressão devido ao custo de gravação no cache.
- O atributo Armazenamento em Cache de Fragmentos Resultantes funciona exclusivamente com arquivos Parquet. Arquivos em outros formatos não são compatíveis.
- Os buffers do atributo Armazenamento em Cache de Fragmentos Resultantes só tentarão armazenar em cache verificações com tamanhos de divisão de arquivos de 128 MB ou mais. Com a configuração padrão do Spark, o Armazenamento em Cache de Fragmentos Resultantes será desabilitado se o tamanho da verificação (tamanho total de todos os arquivos que estão sendo verificados) dividido pelo número de núcleos do executor for menor que 128 MB. Quando qualquer uma das configurações do Spark listadas abaixo for definida, o tamanho da divisão de arquivos será:

  ```
  min(maxPartitionBytes, max(openCostInBytes, scan size / minPartitionNum))
  ```

  - spark.sql. leafNodeDefaultParalelismo (o valor padrão é spark.default.parallelism)
  - spark.sql.files. minPartitionNum (o valor padrão é spark.sql. leafNodeDefaultParalelismo)
  - spark.sql.files. openCostInBytes
  - spark.sql.files. maxPartitionBytes
- O atributo Armazenamento em Cache de Fragmentos Resultantes realiza o armazenamento em cache na granularidade da partição de Resilient Distributed Dataset (RDD - conjunto de dados resilientes e distribuídos). A taxa de redução descrita anteriormente, cujo padrão é 8:1, é avaliada por partição de RDD. Workloads com taxas de redução por RDD maiores e menores que 8:1 podem ter benefícios menores de performance do que workloads com taxas de redução por RDD consistentemente menores que 8:1.
- O atributo Armazenamento em Cache de Fragmentos Resultantes usa um buffer de gravação de 16 MB por padrão para cada partição RDD armazenada em cache. Se mais de 16 MB forem

armazenados em cache por partição RDD, o custo de determinar que uma gravação não é possível pode resultar em uma regressão de desempenho.

- Embora, por padrão, o Armazenamento em Cache de Fragmentos Resultantes não tente armazenar em cache os resultados da partição de RDD com uma taxa de redução menor que 8:1 e limite seu buffer de gravação em 16 MB, esses dois valores podem ser ajustados por meio das seguintes configurações:

```
spark.sql.subResultCache.reductionRatioThreshold (default: 8.0)
spark.sql.subResultCache.maxBufferSize (default: 16MB, max: 64MB)
```

- Vários clusters usando a mesma versão do Amazon EMR podem compartilhar o mesmo local de cache. Para garantir a exatidão dos resultados, o Armazenamento em Cache de Fragmentos Resultantes não usará resultados de cache gravados por diferentes versões do Amazon EMR.
- O cache de fragmentos de resultados será desativado automaticamente para casos de uso do Spark Streaming ou quando RecordServer o Apache Ranger for usado. AWS Lake Formation
- O cache do fragmento resultante read/writes usa buckets EMRFS /S3A e Amazon S3. A criptografia CSE (somente com EMRFS) /SSE S3/SSE KMS é suportada. Para contextualizar, o S3A fornece uma implementação do Hadoop para permitir que um cluster leia e grave dados de e para o Amazon S3. Observe que o suporte para S3A está disponível com o EMR-7.4.0 e superior.

Uso do RAPIDS Accelerator do NVIDIA para Apache Spark

Com o Amazon EMR versão 6.2.0 e posterior, você pode usar o plug-in RAPIDS Accelerator for Apache Spark da Nvidia para acelerar o Spark usando EC2 tipos de instância de unidade de processamento gráfico (GPU). O RAPIDS Accelerator acelerará, por meio da GPU, seus pipelines de ciência de dados do Apache Spark 3.0 sem alterações no código, bem como o processamento de dados e o treinamento de modelos, enquanto reduz substancialmente os custos de infraestrutura.

As seções a seguir orientam você na configuração do cluster do EMR para usar o plug-in Spark-RAPIDS para Spark.

Selecionar tipos de instância

Para usar o plug-in Nvidia Spark-rapids para Spark, os grupos de instâncias principais e de tarefas devem usar tipos de instância de EC2 GPU que atendam aos requisitos de hardware do Spark-rapids. Para ver uma lista completa dos tipos de instâncias de GPU compatíveis com o Amazon EMR, consulte Tipos de instâncias compatíveis no Guia de gerenciamento do Amazon EMR. O tipo

de instância do grupo primário de instâncias pode ser do tipo GPU ou não GPU, mas os tipos de instâncias ARM não são compatíveis.

Definir as configurações da aplicação para o cluster

1. Permitir que o Amazon EMR instale os plug-ins no novo cluster

Para instalar plug-ins, forneça a seguinte configuração ao criar o cluster:

```
{
 "Classification":"spark",
 "Properties":{
  "enableSparkRapids":"true"
 }
}
```

2. Configurar o YARN para usar a GPU

Para obter detalhes sobre como usar a GPU no YARN, consulte Using GPU on YARN na documentação do Apache Hadoop. Os seguintes exemplos são de configurações do YARN para as versões 6.x e 7.x do Amazon EMR:

Amazon EMR 7.x

Exemplo de configuração do YARN para o Amazon EMR 7.x

```
{
    "Classification":"yarn-site",
    "Properties":{
        "yarn.nodemanager.resource-plugins":"yarn.io/gpu",
        "yarn.resource-types":"yarn.io/gpu",
        "yarn.nodemanager.resource-plugins.gpu.allowed-gpu-devices":"auto",
        "yarn.nodemanager.resource-plugins.gpu.path-to-discovery-executables":"/usr/
bin",
        "yarn.nodemanager.linux-container-executor.cgroups.mount":"true",
        "yarn.nodemanager.linux-container-executor.cgroups.mount-path":"/spark-
rapids-cgroup",
        "yarn.nodemanager.linux-container-executor.cgroups.hierarchy":"yarn",
        "yarn.nodemanager.container-
executor.class":"org.apache.hadoop.yarn.server.nodemanager.LinuxContainerExecutor"
    }
},{
    "Classification":"container-executor",
```

```
    "Properties":{

    },
    "Configurations":[
        {
            "Classification":"gpu",
            "Properties":{
                "module.enabled":"true"
            }
        },
        {
            "Classification":"cgroups",
            "Properties":{
                "root":"/spark-rapids-cgroup",
                "yarn-hierarchy":"yarn"
            }
        }
    ]
}
```

Amazon EMR 6.x

Exemplo de configuração do YARN para o Amazon EMR 6.x

```
{
    "Classification":"yarn-site",
    "Properties":{
        "yarn.nodemanager.resource-plugins":"yarn.io/gpu",
        "yarn.resource-types":"yarn.io/gpu",
        "yarn.nodemanager.resource-plugins.gpu.allowed-gpu-devices":"auto",
        "yarn.nodemanager.resource-plugins.gpu.path-to-discovery-executables":"/usr/
bin",
        "yarn.nodemanager.linux-container-executor.cgroups.mount":"true",
        "yarn.nodemanager.linux-container-executor.cgroups.mount-path":"/sys/fs/
cgroup",
        "yarn.nodemanager.linux-container-executor.cgroups.hierarchy":"yarn",
        "yarn.nodemanager.container-
executor.class":"org.apache.hadoop.yarn.server.nodemanager.LinuxContainerExecutor"
    }
},{
    "Classification":"container-executor",
    "Properties":{

    },
```

```
    "Configurations":[
        {
            "Classification":"gpu",
            "Properties":{
                "module.enabled":"true"
            }
        },
        {
            "Classification":"cgroups",
            "Properties":{
                "root":"/sys/fs/cgroup",
                "yarn-hierarchy":"yarn"
            }
        }
    ]
}
```

3. Configurar o Spark para usar o RAPIDS

Veja as configurações necessárias para habilitar o Spark para usar o plugin RAPIDS:

```
{
 "Classification":"spark-defaults",
 "Properties":{
  "spark.plugins":"com.nvidia.spark.SQLPlugin",
  "spark.executor.resource.gpu.discoveryScript":"/usr/lib/spark/scripts/gpu/
getGpusResources.sh",
  "spark.executor.extraLibraryPath":"/usr/local/cuda/targets/x86_64-linux/lib:/usr/
local/cuda/extras/CUPTI/lib64:/usr/local/cuda/compat/lib:/usr/local/cuda/lib:/usr/
local/cuda/lib64:/usr/lib/hadoop/lib/native:/usr/lib/hadoop-lzo/lib/native:/docker/usr/
lib/hadoop/lib/native:/docker/usr/lib/hadoop-lzo/lib/native"
 }
}
```

XGBoost4A biblioteca J-Spark na XGBoost documentação também está disponível quando o plug-in Spark RAPIDS está habilitado em seu cluster. Você pode usar a seguinte configuração para se integrar XGBoost ao seu trabalho do Spark:

```
{
 "Classification":"spark-defaults",
 "Properties":{
  "spark.submit.pyFiles":"/usr/lib/spark/jars/xgboost4j-spark_3.0-1.4.2-0.3.0.jar"
```

```
 }
}
```

Para conhecer configurações adicionais do Spark que você pode usar para ajustar um cluster do EMR acelerado por GPU, consulte o Guia de ajuste do acelerador Rapids para Apache Spark na documentação do Nvidia.github.io.

4. Configurar o YARN Capacity Scheduler

`DominantResourceCalculator` deve estar configurado para permitir a programação e o isolamento da GPU. Para obter mais informações, consulte Usar GPU no YARN na documentação do Apache Hadoop.

```
{
 "Classification":"capacity-scheduler",
 "Properties":{
  "yarn.scheduler.capacity.resource-
calculator":"org.apache.hadoop.yarn.util.resource.DominantResourceCalculator"
 }
}
```

5. Criação de um arquivo JSON para incluir as configurações

Você pode criar um arquivo JSON que contenha sua configuração para usar o plug-in RAPIDS no cluster do Spark. Você fornece o arquivo posteriormente ao executar o cluster.

Você pode armazenar o arquivo localmente ou no S3. Para obter mais informações sobre como fornecer configurações de aplicações para os clusters, consulte Configurar aplicações.

Use os exemplos de arquivos a seguir como modelos para criar suas próprias configurações.

Amazon EMR 7.x

Arquivo de exemplo **`my-configurations.json`** para o Amazon EMR 7.x

```
[
    {
        "Classification":"spark",
        "Properties":{
            "enableSparkRapids":"true"
        }
    },
    {
```

```
        "Classification":"yarn-site",
        "Properties":{
            "yarn.nodemanager.resource-plugins":"yarn.io/gpu",
            "yarn.resource-types":"yarn.io/gpu",
            "yarn.nodemanager.resource-plugins.gpu.allowed-gpu-devices":"auto",
            "yarn.nodemanager.resource-plugins.gpu.path-to-discovery-executables":"/
usr/bin",
            "yarn.nodemanager.linux-container-executor.cgroups.mount":"true",
            "yarn.nodemanager.linux-container-executor.cgroups.mount-path":"/spark-
rapids-cgroup",
            "yarn.nodemanager.linux-container-executor.cgroups.hierarchy":"yarn",
            "yarn.nodemanager.container-
executor.class":"org.apache.hadoop.yarn.server.nodemanager.LinuxContainerExecutor"
        }
    },
    {
        "Classification":"container-executor",
        "Properties":{

        },
        "Configurations":[
            {
                "Classification":"gpu",
                "Properties":{
                    "module.enabled":"true"
                }
            },
            {
                "Classification":"cgroups",
                "Properties":{
                    "root":"/spark-rapids-cgroup",
                    "yarn-hierarchy":"yarn"
                }
            }
        ]
    },
    {
        "Classification":"spark-defaults",
        "Properties":{
            "spark.plugins":"com.nvidia.spark.SQLPlugin",
            "spark.executor.resource.gpu.discoveryScript":"/usr/lib/spark/scripts/
gpu/getGpusResources.sh",
            "spark.executor.extraLibraryPath":"/usr/local/cuda/targets/x86_64-
linux/lib:/usr/local/cuda/extras/CUPTI/lib64:/usr/local/cuda/compat/lib:/usr/local/
```

```
cuda/lib:/usr/local/cuda/lib64:/usr/lib/hadoop/lib/native:/usr/lib/hadoop-lzo/lib/
native:/docker/usr/lib/hadoop/lib/native:/docker/usr/lib/hadoop-lzo/lib/native",
            "spark.submit.pyFiles":"/usr/lib/spark/jars/xgboost4j-
spark_3.0-1.4.2-0.3.0.jar",
            "spark.rapids.sql.concurrentGpuTasks":"1",
            "spark.executor.resource.gpu.amount":"1",
            "spark.executor.cores":"2",
            "spark.task.cpus":"1",
            "spark.task.resource.gpu.amount":"0.5",
            "spark.rapids.memory.pinnedPool.size":"0",
            "spark.executor.memoryOverhead":"2G",
            "spark.locality.wait":"0s",
            "spark.sql.shuffle.partitions":"200",
            "spark.sql.files.maxPartitionBytes":"512m"
        }
    },
    {
        "Classification":"capacity-scheduler",
        "Properties":{
            "yarn.scheduler.capacity.resource-
calculator":"org.apache.hadoop.yarn.util.resource.DominantResourceCalculator"
        }
    }
]
```

Amazon EMR 6.x

Arquivo de exemplo **`my-configurations.json`** para o Amazon EMR 6.x

```
[
    {
        "Classification":"spark",
        "Properties":{
            "enableSparkRapids":"true"
        }
    },
    {
        "Classification":"yarn-site",
        "Properties":{
            "yarn.nodemanager.resource-plugins":"yarn.io/gpu",
            "yarn.resource-types":"yarn.io/gpu",
            "yarn.nodemanager.resource-plugins.gpu.allowed-gpu-devices":"auto",
            "yarn.nodemanager.resource-plugins.gpu.path-to-discovery-executables":"/
usr/bin",
```

```
            "yarn.nodemanager.linux-container-executor.cgroups.mount":"true",
            "yarn.nodemanager.linux-container-executor.cgroups.mount-path":"/sys/fs/
cgroup",
            "yarn.nodemanager.linux-container-executor.cgroups.hierarchy":"yarn",
            "yarn.nodemanager.container-
executor.class":"org.apache.hadoop.yarn.server.nodemanager.LinuxContainerExecutor"
        }
    },
    {
        "Classification":"container-executor",
        "Properties":{

        },
        "Configurations":[
            {
                "Classification":"gpu",
                "Properties":{
                    "module.enabled":"true"
                }
            },
            {
                "Classification":"cgroups",
                "Properties":{
                    "root":"/sys/fs/cgroup",
                    "yarn-hierarchy":"yarn"
                }
            }
        ]
    },
    {
        "Classification":"spark-defaults",
        "Properties":{
            "spark.plugins":"com.nvidia.spark.SQLPlugin",
            "spark.executor.resource.gpu.discoveryScript":"/usr/lib/spark/scripts/
gpu/getGpusResources.sh",
            "spark.executor.extraLibraryPath":"/usr/local/cuda/targets/x86_64-
linux/lib:/usr/local/cuda/extras/CUPTI/lib64:/usr/local/cuda/compat/lib:/usr/local/
cuda/lib:/usr/local/cuda/lib64:/usr/lib/hadoop/lib/native:/usr/lib/hadoop-lzo/lib/
native:/docker/usr/lib/hadoop/lib/native:/docker/usr/lib/hadoop-lzo/lib/native",
            "spark.submit.pyFiles":"/usr/lib/spark/jars/xgboost4j-
spark_3.0-1.4.2-0.3.0.jar",
            "spark.rapids.sql.concurrentGpuTasks":"1",
            "spark.executor.resource.gpu.amount":"1",
            "spark.executor.cores":"2",
```

```
            "spark.task.cpus":"1",
            "spark.task.resource.gpu.amount":"0.5",
            "spark.rapids.memory.pinnedPool.size":"0",
            "spark.executor.memoryOverhead":"2G",
            "spark.locality.wait":"0s",
            "spark.sql.shuffle.partitions":"200",
            "spark.sql.files.maxPartitionBytes":"512m"
        }
    },
    {
        "Classification":"capacity-scheduler",
        "Properties":{
            "yarn.scheduler.capacity.resource-
calculator":"org.apache.hadoop.yarn.util.resource.DominantResourceCalculator"
        }
    }
]
```

Adicionar uma ação de bootstrap para o cluster

Para obter mais informações sobre como fornecer scripts de ação de bootstrap ao criar o cluster, consulte Bootstrap action basics no Guia de gerenciamento do Amazon EMR.

Os seguintes exemplos de scripts mostram como criar um arquivo de ação de bootstrap para o Amazon EMR 6.x e 7.x:

Amazon EMR 7.x

Arquivo de exemplo **`my-bootstrap-action.sh`** para o Amazon EMR 7.x

Para usar o YARN no gerenciamento de recursos de GPU com as versões 7.x do Amazon EMR, você deve montar o CGroup v1 manualmente no seu cluster. Isso pode ser feito com um script de ação de bootstrap, conforme mostrado neste exemplo.

```
#!/bin/bash
set -ex

sudo mkdir -p /spark-rapids-cgroup/devices
sudo mount -t cgroup -o devices cgroupv1-devices /spark-rapids-cgroup/devices
sudo chmod a+rwx -R /spark-rapids-cgroup
```

Amazon EMR 6.x

Arquivo de exemplo **`my-bootstrap-action.sh`** para o Amazon EMR 6.x

Nas versões 6.x do Amazon EMR, você deve abrir as permissões do CGroup para o YARN no seu cluster. Isso pode ser feito com um script de ação de bootstrap, conforme mostrado neste exemplo.

```
#!/bin/bash
set -ex

sudo chmod a+rwx -R /sys/fs/cgroup/cpu,cpuacct
sudo chmod a+rwx -R /sys/fs/cgroup/devices
```

Iniciar o cluster

A última etapa é iniciar o cluster com as configurações de cluster mencionadas acima. Veja um exemplo de comando para executar um cluster pela CLI do Amazon EMR:

```
 aws emr create-cluster \
--release-label emr-7.9.0 \
--applications Name=Hadoop Name=Spark \
--service-role EMR_DefaultRole_V2 \
--ec2-attributes KeyName=my-key-pair,InstanceProfile=EMR_EC2_DefaultRole \
--instance-groups InstanceGroupType=MASTER,InstanceCount=1,InstanceType=m4.4xlarge \

                  InstanceGroupType=CORE,InstanceCount=1,InstanceType=g4dn.2xlarge \

                  InstanceGroupType=TASK,InstanceCount=1,InstanceType=g4dn.2xlarge \
--configurations file:///my-configurations.json \
--bootstrap-actions Name='My Spark Rapids Bootstrap action',Path=s3://amzn-s3-demo-
bucket/my-bootstrap-action.sh
```

Acessar o shell do Spark

O shell do Spark é baseado no Scala REPL (Read-Eval-Print-Loop). Ele permite criar programas Spark interativamente e enviar trabalhos à estrutura. Você pode acessar o shell do Spark conectando-se ao nó primário com SSH e invocando `spark-shell`. Para obter informações sobre a conexão ao nó primário, consulte Connect to the primary node using SSH no Guia de

gerenciamento do Amazon EMR. Os exemplos a seguir usam logs de acesso do Servidor HTTP Apache armazenados no Amazon S3.

> **Note**
>
> O bucket nesses exemplos está disponível para clientes que podem acessar a região Leste dos EUA (Norte da Virgínia).

Por padrão, o shell do Spark cria seu próprio SparkContextobjeto chamado`sc`. Você pode usar esse contexto se for necessário no REPL. sqlContexttambém está disponível na concha e é uma HiveContext.

Example Usar o shell do Spark para contar as ocorrências de uma string em um arquivo armazenado no Amazon S3

Este exemplo usa `sc` para ler um arquivo de texto armazenado no Amazon S3.

```
scala> sc
res0: org.apache.spark.SparkContext = org.apache.spark.SparkContext@404721db

scala> val textFile = sc.textFile("s3://elasticmapreduce/samples/hive-ads/tables/
impressions/dt=2009-04-13-08-05/ec2-0-51-75-39.amazon.com-2009-04-13-08-05.log")
```

O Spark cria o textFile e a estrutura de dados associada. Em seguida, o exemplo conta o número de linhas no arquivo de log com a string "cartoonnetwork.com":

```
scala> val linesWithCartoonNetwork = textFile.filter(line =>
 line.contains("cartoonnetwork.com")).count()
linesWithCartoonNetwork: org.apache.spark.rdd.RDD[String] = MapPartitionsRDD[2] at
 filter at <console>:23
<snip>
<Spark program runs>
scala> linesWithCartoonNetwork
res2: Long = 9
```

Example Usar o shell do Spark baseado em Python para contar as ocorrências de uma string em um arquivo armazenado no Amazon S3

O Spark também inclui um shell baseado em Python, `pyspark`, que você pode usar para gerar protótipos de programas Spark escritos em Python. Assim como com`spark-shell`, invoque `pyspark` no nó primário; ele também tem o mesmo SparkContextobjeto.

```
>>> sc
<pyspark.context.SparkContext object at 0x7fe7e659fa50>
>>> textfile = sc.textFile("s3://elasticmapreduce/samples/hive-ads/tables/impressions/
dt=2009-04-13-08-05/ec2-0-51-75-39.amazon.com-2009-04-13-08-05.log")
```

O Spark cria o textFile e a estrutura de dados associada. Em seguida, o exemplo conta o número de linhas no arquivo de log com a string "cartoonnetwork.com".

```
>>> linesWithCartoonNetwork = textfile.filter(lambda line: "cartoonnetwork.com" in
 line).count()
15/06/04 17:12:22 INFO lzo.GPLNativeCodeLoader: Loaded native gpl library from the
 embedded binaries
15/06/04 17:12:22 INFO lzo.LzoCodec: Successfully loaded & initialized native-lzo
 library [hadoop-lzo rev EXAMPLE]
15/06/04 17:12:23 INFO fs.EmrFileSystem: Consistency disabled, using
 com.amazon.ws.emr.hadoop.fs.s3n.S3NativeFileSystem as filesystem implementation
<snip>
<Spark program continues>
>>> linesWithCartoonNetwork
9
```

Use o Amazon SageMaker Spark para aprendizado de máquina

Quando você usa as versões 5.11.0 e posteriores do Amazon EMR, o componente `aws-sagemaker-spark-sdk` é instalado com o Spark. Esse componente instala o Amazon SageMaker Spark e as dependências associadas para a integração do Spark com a Amazon. SageMaker Você pode usar o Amazon SageMaker Spark para criar pipelines de aprendizado de máquina (ML) do Spark usando os estágios da Amazon. SageMaker Para obter mais informações, consulte o README do Amazon SageMaker Spark sobre GitHub e como usar o Apache Spark com a Amazon SageMaker no Amazon Developer Guide. SageMaker

Escrever uma aplicação do Spark

Aplicativos Spark podem ser escritos em Scala, Java ou Python. Existem vários exemplos de aplicações do Spark no tópico Spark examples da documentação do Apache Spark. O exemplo "Estimating Pi" é mostrado abaixo em três aplicativos com suporte nativo. Você também pode ver exemplos completos em `$SPARK_HOME/examples` e em GitHub. Para obter mais informações sobre como criar JARs para o Spark, consulte o tópico Início rápido na documentação do Apache Spark.

Scala

Para evitar problemas de compatibilidade com o Scala, sugerimos que você use as dependências do Spark para a versão correta do Scala ao compilar uma aplicação do Spark para um cluster do Amazon EMR. A versão do Scala que você deve usar depende da versão do Spark instalada no cluster. Por exemplo, a versão 5.30.1 do Amazon EMR usa o Spark 2.4.5, que é criado com o Scala 2.11. Se o cluster usa a versão 5.30.1 do Amazon EMR, use as dependências do Spark para o Scala 2.11. Para obter mais informações sobre as versões do Scala usadas pelo Spark, consulte a documentação do Apache Spark.

```
package org.apache.spark.examples
import scala.math.random
import org.apache.spark._

/** Computes an approximation to pi */
object SparkPi {
  def main(args: Array[String]) {
    val conf = new SparkConf().setAppName("Spark Pi")
    val spark = new SparkContext(conf)
    val slices = if (args.length > 0) args(0).toInt else 2
    val n = math.min(100000L * slices, Int.MaxValue).toInt // avoid overflow
    val count = spark.parallelize(1 until n, slices).map { i =>
      val x = random * 2 - 1
      val y = random * 2 - 1
      if (x*x + y*y < 1) 1 else 0
    }.reduce(_ + _)
    println("Pi is roughly " + 4.0 * count / n)
    spark.stop()
  }
}
```

Java

```
package org.apache.spark.examples;

import org.apache.spark.SparkConf;
import org.apache.spark.api.java.JavaRDD;
import org.apache.spark.api.java.JavaSparkContext;
import org.apache.spark.api.java.function.Function;
import org.apache.spark.api.java.function.Function2;

import java.util.ArrayList;
import java.util.List;

/**
 * Computes an approximation to pi
 * Usage: JavaSparkPi [slices]
 */
public final class JavaSparkPi {

  public static void main(String[] args) throws Exception {
    SparkConf sparkConf = new SparkConf().setAppName("JavaSparkPi");
    JavaSparkContext jsc = new JavaSparkContext(sparkConf);

    int slices = (args.length == 1) ? Integer.parseInt(args[0]) : 2;
    int n = 100000 * slices;
    List<Integer> l = new ArrayList<Integer>(n);
    for (int i = 0; i < n; i++) {
      l.add(i);
    }

    JavaRDD<Integer> dataSet = jsc.parallelize(l, slices);

    int count = dataSet.map(new Function<Integer, Integer>() {
      @Override
      public Integer call(Integer integer) {
        double x = Math.random() * 2 - 1;
        double y = Math.random() * 2 - 1;
        return (x * x + y * y < 1) ? 1 : 0;
      }
    }).reduce(new Function2<Integer, Integer, Integer>() {
      @Override
      public Integer call(Integer integer, Integer integer2) {
        return integer + integer2;
```

```
      }
    });

    System.out.println("Pi is roughly " + 4.0 * count / n);

    jsc.stop();
  }
}
```

Python

```
import argparse
import logging
from operator import add
from random import random

from pyspark.sql import SparkSession

logger = logging.getLogger(__name__)
logging.basicConfig(level=logging.INFO, format="%(levelname)s: %(message)s")


def calculate_pi(partitions, output_uri):
    """
    Calculates pi by testing a large number of random numbers against a unit circle
    inscribed inside a square. The trials are partitioned so they can be run in
    parallel on cluster instances.

    :param partitions: The number of partitions to use for the calculation.
    :param output_uri: The URI where the output is written, typically an Amazon S3
                       bucket, such as 's3://example-bucket/pi-calc'.
    """

    def calculate_hit(_):
        x = random() * 2 - 1
        y = random() * 2 - 1
        return 1 if x**2 + y**2 < 1 else 0

    tries = 100000 * partitions
    logger.info(
        "Calculating pi with a total of %s tries in %s partitions.", tries, partitions
    )
```

```
    with SparkSession.builder.appName("My PyPi").getOrCreate() as spark:
        hits = (
            spark.sparkContext.parallelize(range(tries), partitions)
            .map(calculate_hit)
            .reduce(add)
        )
        pi = 4.0 * hits / tries
        logger.info("%s tries and %s hits gives pi estimate of %s.", tries, hits, pi)
        if output_uri is not None:
            df = spark.createDataFrame([(tries, hits, pi)], ["tries", "hits", "pi"])
            df.write.mode("overwrite").json(output_uri)


if __name__ == "__main__":
    parser = argparse.ArgumentParser()
    parser.add_argument(
        "--partitions",
        default=2,
        type=int,
        help="The number of parallel partitions to use when calculating pi.",
    )
    parser.add_argument(
        "--output_uri", help="The URI where output is saved, typically an S3 bucket."
    )
    args = parser.parse_args()

    calculate_pi(args.partitions, args.output_uri)
```

Melhorar a performance do Spark com Amazon S3

O Amazon EMR oferece atributos para ajudar na otimização da performance ao usar o Spark para consultar, ler e gravar dados salvos no Amazon S3.

O S3 Select pode melhorar a performance de consultas para arquivos CSV e JSON em algumas aplicações ao enviar o processamento para o Amazon S3.

O committer otimizado para EMRFS S3 é uma alternativa à OutputCommitterclasse, que usa o recurso de uploads de várias partes do EMRFS para melhorar o desempenho ao gravar arquivos Parquet no Amazon S3 usando Spark e conjuntos de dados. DataFrames

Tópicos

- Usar o S3 Select com Spark para melhorar a performance das consultas
- Usar o confirmador otimizado para EMRFS S3
- Use o protocolo de confirmação otimizada para EMRFS S3
- Tentar novamente solicitações do Amazon S3 com EMRFS

Usar o S3 Select com Spark para melhorar a performance das consultas

> **Important**
>
> O Amazon S3 Select não está mais disponível para novos clientes. Os clientes existentes do Amazon S3 Select podem continuar usando o recurso normalmente. Saiba mais

Com as versões 5.17.0 e posteriores do Amazon EMR, você pode usar o S3 Select com Spark no Amazon EMR. O S3 Select possibilita que as aplicações recuperem apenas um subconjunto dos dados de um objeto. Para o Amazon EMR, o trabalho computacional de filtrar grandes conjuntos de dados para processamento é enviado do cluster para o Amazon S3, o que pode melhorar a performance em algumas aplicações e reduz o volume de dados transferidos entre o Amazon EMR e o Amazon S3.

O S3 Select é compatível com arquivos CSV e JSON usando os valores `s3selectCSV` e `s3selectJSON` para especificar o formato de dados. Para ter mais informações e exemplos, consulte Especificar o S3 Select no seu código.

O S3 Select é adequado para minha aplicação?

Recomendamos que você avalie seus aplicativos com e sem o S3 Selecione para ver se o uso pode ser adequado para o aplicativo.

Use as seguintes diretrizes para determinar se o seu aplicativo é adequado para o uso do S3 Select:

- Sua consulta filtra mais de metade do conjunto de dados original.
- Sua conexão de rede entre o Amazon S3 e o cluster do Amazon EMR tem boa velocidade de transferência e largura de banda disponível. O Amazon S3 não compacta respostas HTTP. Portanto, é provável que o tamanho da resposta aumente para arquivos de entrada compactados.

Considerações e limitações

- A criptografia do lado do servidor do Amazon S3 com chaves de criptografia fornecidas pelo cliente (SSE-C) e a criptografia do lado do cliente não são compatíveis.
- A propriedade `AllowQuotedRecordDelimiters` não é compatível. Se essa propriedade for especificada, a consulta falhará.
- Somente arquivos CSV e JSON no formato UTF-8 são compatíveis. Não CSVs há suporte para várias linhas.
- Somente arquivos descompactados ou gzip são compatíveis.
- As opções CSV e JSON do Spark, como `nanValue`, `positiveInf`, `negativeInf` e opções relacionadas a registros corrompidos (por exemplo, modo dropmalformed e failfast) não são compatíveis.
- O uso de vírgulas (,) em casas decimais não é compatível. Por exemplo, `10,000` não é compatível, mas `10000` é.
- Caracteres de comentário na última linha não são compatíveis.
- Linhas vazias no final de um arquivo não são processadas.
- Os seguintes filtros não são enviados para o Amazon S3:
  - Funções agregadas, como `COUNT()` e `SUM()`.
  - Filtros que `CAST()` um atributo. Por exemplo, `.CAST(stringColumn as INT) = 1`
  - Filtros com um atributo que é um objeto ou complexo. Por exemplo, `.intArray[1] = 1, objectColumn.objectNumber = 1`
  - Filtros para os quais o valor não é um valor literal. Por exemplo, `intColumn1 = intColumn2`.
  - Somente tipos de dados compatíveis com o S3 Select são compatíveis com as limitações documentadas.

Especificar o S3 Select no seu código

Os exemplos a seguir demonstram como especificar o S3 Select para CSV usando Scala, SQL, R e. PySpark Você pode usar o S3 Select para JSON da mesma forma. Para obter uma lista de opções, os valores padrão e limitações, consulte Opções.

PySpark

```
spark
  .read
```

```
  .format("s3selectCSV") // "s3selectJson" for Json
  .schema(...) // optional, but recommended
  .options(...) // optional
  .load("s3://path/to/my/datafiles")
```

R

```
read.df("s3://path/to/my/datafiles", "s3selectCSV", schema, header = "true",
 delimiter = "\t")
```

Scala

```
spark
  .read
  .format("s3selectCSV") // "s3selectJson" for Json
  .schema(...) // optional, but recommended
  .options(...) // optional. Examples:
  // .options(Map("quote" -> "\'", "header" -> "true")) or
  // .option("quote", "\'").option("header", "true")
  .load("s3://path/to/my/datafiles")
```

SQL

```
CREATE TEMPORARY VIEW MyView (number INT, name STRING) USING s3selectCSV OPTIONS
 (path "s3://path/to/my/datafiles", header "true", delimiter "\t")
```

Opções

As seguintes opções estão disponíveis ao usar `s3selectCSV` e `s3selectJSON`. Se não for especificado, os valores padrão serão usados.

Opções com o S3selectCSV

Opção	Padrão	Uso
`compression`	`"none"`	Indica se a compactação é usada. `"gzip"` é a única configuração compatível além de `"none"`.

Opção	Padrão	Uso
`delimiter`	`","`	Especifica o delimitador de campo.
`quote`	`'\"'`	Especifica o caractere de aspas. Não há suporte para a especificação de uma string vazia e isso resulta em um erro de XML malformado.
`escape`	`'\\'`	Especifica o caractere de escape.
`header`	`"false"`	`"false"` especifica que não há cabeçalho. `"true"` especifica que o cabeçalho está na primeira linha. Somente cabeçalhos na primeira linha são suportados e linhas vazias antes de um cabeçalho não são compatíveis.
comment	`"#"`	Especifica o caractere de comentário. O indicador de comentários não pode ser desativado. Em outras palavras, um valor de `\u0000` não é suportado.
`nullValue`	`""`	

Opções com S3selectJSON

Opção	Padrão	Uso
`compression`	`"none"`	Indica se a compactação é usada. `"gzip"` é a única configuração compatível além de `"none"`.
`multiline`	"falso"	`"false"` especifica que o JSON está no formato `LINES` do S3 Select, o que significa que cada linha nos dados de entrada contém um único objeto JSON. `"true"` especifica que o JSON está no formato `DOCUMENT` do S3 Select, o que significa que um objeto JSON podem abranger várias linhas nos dados de entrada.

Usar o confirmador otimizado para EMRFS S3

O committer otimizado para EMRFS S3 é uma OutputCommitterimplementação alternativa otimizada para gravar arquivos no Amazon S3 ao usar o EMRFS. O confirmador otimizado para EMRFS S3 melhora a performance da aplicação ao evitar operações de listar e renomear no Amazon S3 durante o trabalho e fases de confirmação da tarefa. O confirmador está disponível com as versões 5.19.0 e posteriores do Amazon EMR e é habilitado por padrão nas versões 5.20.0 e posteriores do Amazon EMR. O committer é usado para trabalhos do Spark que usam Spark ou conjuntos de dados DataFrames. Desde o Amazon EMR 6.4.0, esse confirmador pode ser usado para todos os formatos comuns, incluindo parquet, ORC e formatos baseados em texto (incluindo CSV e JSON). Para versões anteriores ao Amazon EMR 6.4.0, somente o formato Parquet é compatível. Há circunstâncias em que o committer não é usado. Para obter mais informações, consulte Requisitos do confirmador otimizado para EMRFS S3.

Tópicos

- Requisitos do confirmador otimizado para EMRFS S3
- O confirmador otimizado para EMRFS S3 e carregamentos multipart
- Considerações sobre ajuste de tarefas
- Ative o confirmador otimizado para EMRFS S3 para o Amazon EMR 5.19.0

Requisitos do confirmador otimizado para EMRFS S3

O committer otimizado para EMRFS S3 é usado quando as seguintes condições são atendidas:

- Você executa trabalhos do Spark que usam o Spark ou conjuntos de dados para gravar arquivos no Amazon S3. DataFrames Desde o Amazon EMR 6.4.0, esse confirmador pode ser usado para todos os formatos comuns, incluindo parquet, ORC e formatos baseados em texto (incluindo CSV e JSON). Para versões anteriores ao Amazon EMR 6.4.0, somente o formato Parquet é compatível.
- Carregamentos multipart estão habilitados no Amazon EMR. Esse é o padrão. Para obter mais informações, consulte O confirmador otimizado para EMRFS S3 e carregamentos multipart.
- O suporte a formato de arquivo integrado do Spark é usado. O suporte a formato de arquivo integrado é usado nas seguintes circunstâncias:
  - Para tabelas do Hive Metastore, quando `spark.sql.hive.convertMetastoreParquet` é definido como `true` para tabelas do Parquet ou `spark.sql.hive.convertMetastoreOrc` é definido como `true` para tabelas do Orc com as versões 6.4.0 ou superiores do Amazon EMR. Essas são as configurações padrão.
  - Quando os trabalhos são gravados em fontes de dados ou tabelas do Parquet, por exemplo, a tabela de destino é criada com a cláusula `USING parquet`.
  - Quando os trabalhos gravam em tabelas Parquet de metastore do Hive não particionadas. O suporte ao Parquet incorporado do Spark oferece suporte a tabelas particionadas do Hive, o que é uma limitação conhecida. Para obter mais informações, consulte Conversão da tabela Parquet do Hive metastore no Apache Spark e no Guia de conjuntos de dados. DataFrames
- As operações de trabalhos do Spark que gravam em um local de partição padrão, como `${table_location}/k1=v1/k2=v2/`, usam o confirmador. O confirmador não será usado se uma operação de trabalho gravar em um local de partição personalizado, por exemplo, se o local de uma partição personalizado for definido usando o comando `ALTER TABLE SQL`.
- Os valores a seguir para o Spark devem ser usados:
  - A propriedade `spark.sql.parquet.fs.optimized.committer.optimization-enabled` deve ser definida como `true`. Essa é a configuração padrão com as versões 5.20.0 e

posteriores do Amazon EMR. Com o Amazon EMR 5.19.0, o valor padrão é `false`. Para obter informações sobre como configurar esse valor, consulte Ative o confirmador otimizado para EMRFS S3 para o Amazon EMR 5.19.0.

- Se você está gravando em tabelas da metastore do Hive não particionadas, somente os formatos de arquivo Parquet e Orc são compatíveis. `spark.sql.hive.convertMetastoreParquet` deve ser definido como `true` se você está gravando em tabelas da metastore do Hive não particionadas em Paquet. `spark.sql.hive.convertMetastoreOrc` deve ser definido como `true` se estiver gravando em tabelas da metastore do Hive não particionadas em Orc. Essas são as configurações padrão.
- `spark.sql.parquet.output.committer.class` deve ser definido como `com.amazon.emr.committer.EmrOptimizedSparkSqlParquetOutputCommitter`. Essa é a configuração padrão.
- `spark.sql.sources.commitProtocolClass` deve ser definido como `org.apache.spark.sql.execution.datasources.SQLEmrOptimizedCommitProtocol` ou `org.apache.spark.sql.execution.datasources.SQLHadoopMapReduceCommitProtocol`. `org.apache.spark.sql.execution.datasources.SQLEmrOptimizedCommitProtocol` é a configuração padrão para as versões 5.30.0 e superiores do Amazon EMR série 5.x e para as versões 6.2.0 e superiores do Amazon EMR série 6.x. `org.apache.spark.sql.execution.datasources.SQLHadoopMapReduceCommitProtocol` é a configuração padrão para as versões anteriores do Amazon EMR.
- Se os trabalhos do Spark substituírem os conjuntos de dados Parquet por colunas de partição dinâmica, as opções de gravação `partitionOverwriteMode` e `spark.sql.sources.partitionOverwriteMode` deverão ser definidas como `static`. Essa é a configuração padrão.

> **Note**
>
> A opção de gravação `partitionOverwriteMode` foi introduzida no Spark 2.4.0. Para a versão 2.3.2 do Spark, incluída com a versão 5.19.0 do Amazon EMR, defina a propriedade `spark.sql.sources.partitionOverwriteMode`.

Ocasiões em que o confirmador otimizado para EMRFS S3 não é usado

Geralmente, o confirmador otimizado para EMRFS S3 não é usado nas situações a seguir.

Situação	Por que o confirmador não é usado
Quando você grava no HDFS	O confirmador só permite gravação no Amazon S3 com o uso do EMRFS.
Quando você usa o sistema de arquivos S3A	O confirmador só é compatível com o EMRFS.
Quando você usa MapReduce nossa API RDD do Spark	O committer só oferece suporte ao uso de SparkSQL ou Dataset DataFrame. APIs

Os exemplos do Scala a seguir demonstram algumas situações adicionais que impedem que o confirmador otimizado para EMRFS S3 seja usado por completo (o primeiro exemplo) e em parte (o segundo exemplo).

Example – Modo de substituição de partição dinâmica

O exemplo do Scala a seguir instrui o Spark a usar um algoritmo de confirmação diferente, o que impede totalmente o uso do confirmador otimizado para EMRFS S3. O código define a propriedade `partitionOverwriteMode` como `dynamic` para subtituir somente as partições nas quais você está gravando dados. Em seguida, as colunas de partição dinâmica são especificadas por `partitionBy` e o modo de gravação é definido como `overwrite`.

```
val dataset = spark.range(0, 10)
  .withColumn("dt", expr("date_sub(current_date(), id)"))

dataset.write.mode("overwrite")
  .option("partitionOverwriteMode", "dynamic")
  .partitionBy("dt")
  .parquet("s3://amzn-s3-demo-bucket1/output")
```

Você deve definir todas as três configurações para evitar o uso do confirmador otimizado para EMRFS S3. Quando você faz isso, o Spark executa um algoritmo de confirmação diferente, especificado no protocolo de confirmação do Spark. Para versões do Amazon EMR 5.x anteriores à 5.30.0 e para versões do Amazon EMR 6.x anteriores à 6.2.0, o protocolo de confirmação usa o diretório de teste do Spark, que é um diretório temporário criado no local de saída que começa com

`.spark-staging`. O algoritmo renomeia sequencialmente diretórios de partição, o que pode afetar negativamente a performance. Para obter mais informações sobre as versões 5.30.0 e posteriores e 6.2.0 e posteriores do Amazon EMR, consulte. Use o protocolo de confirmação otimizada para EMRFS S3

O algoritmo no Spark 2.4.0 segue estas etapas:

1. As tentativas de tarefa gravam a saída nos diretórios de partição do diretório de teste do Spark, por exemplo, `${outputLocation}/spark-staging-${jobID}/k1=v1/k2=v2/`.
2. Para cada partição gravada, a tentativa de tarefa acompanha os caminhos de partição relativos, por exemplo, `k1=v1/k2=v2`.
3. Quando uma tarefa é concluída com êxito, ela fornece o driver com todos os caminhos de partição relativos que ela controlou.
4. Depois que todas as tarefas forem concluídas, a fase de confirmação do trabalho coletará todos os diretórios da partição que as tentativas de tarefas bem-sucedidas gravaram no diretório de preparação do Spark. O Spark renomeia sequencialmente cada um desses diretórios para o local de saída final usando as operações para renomear a árvore de diretórios.
5. O diretório de preparação é excluído antes de a fase de confirmação de trabalho ser concluída.

Example – Local de partição personalizado

Neste exemplo, o código Scala insere em duas partições. Uma partição tem um local de partição personalizado. A outra partição usa o local de partição padrão. O committer otimizado para EMRFS S3 é usado somente para gravar a saída da tarefa na partição que usa o local de partição padrão.

```
val table = "dataset"
val location = "s3://bucket/table"

spark.sql(s"""
  CREATE TABLE $table (id bigint, dt date)
  USING PARQUET PARTITIONED BY (dt)
  LOCATION '$location'
""")

// Add a partition using a custom location
val customPartitionLocation = "s3://bucket/custom"
spark.sql(s"""
  ALTER TABLE $table ADD PARTITION (dt='2019-01-28')
```

```
  LOCATION '$customPartitionLocation'
""")

// Add another partition using default location
spark.sql(s"ALTER TABLE $table ADD PARTITION (dt='2019-01-29')")

def asDate(text: String) = lit(text).cast("date")

spark.range(0, 10)
  .withColumn("dt",
    when($"id" > 4, asDate("2019-01-28")).otherwise(asDate("2019-01-29")))
  .write.insertInto(table)
```

O código Scala cria os seguintes objetos do Amazon S3:

```
custom/part-00001-035a2a9c-4a09-4917-8819-e77134342402.c000.snappy.parquet
custom_$folder$
table/_SUCCESS
table/dt=2019-01-29/part-00000-035a2a9c-4a09-4917-8819-e77134342402.c000.snappy.parquet
table/dt=2019-01-29_$folder$
table_$folder$
```

Ao gravar em partições em locais personalizados, o Spark usa um algoritmo de confirmação semelhante ao exemplo anterior, que é descrito abaixo. Como no exemplo anterior, o algoritmo resulta em renomeações sequenciais, o que pode afetar negativamente a performance.

1. Ao gravar a saída em uma partição em um local personalizado, as tarefas gravam em um arquivo no diretório de preparação do Spark, que é criado no local de saída final. O nome do arquivo inclui um UUID aleatório para evitar colisões de nomes de arquivos. A tentativa de tarefa controla cada arquivo junto com o caminho de saída final desejado.
2. Quando uma tarefa é concluída com êxito, ela fornece o driver com os arquivos e os caminhos desejados de saída final.
3. Depois que todas as tarefas forem concluídas, a fase de confirmação do trabalho renomeará sequencialmente todos os arquivos que foram gravados para partições em locais personalizados em seus caminhos de saída final.
4. O diretório de preparação é excluído antes de a fase de confirmação de trabalho ser concluída.

O confirmador otimizado para EMRFS S3 e carregamentos multipart

Para usar o confirmador otimizado para EMRFS S3, uploads de várias partes devem estar habilitados no Amazon EMR. Multipart uploads são habilitados por padrão. Você pode habilitá-los novamente, se necessário. Para obter mais informações, consulte Configure multipart upload for Amazon S3 (Configurar o carregamento fracionado no Amazon S3) no Guia de gerenciamento do Amazon EMR.

O EMRFS S3 Optimized Committer usa as características semelhantes a transações de multipart uploads para garantir que os arquivos gravados por tentativas de tarefas aparecem apenas no local de saída do trabalho após a confirmação da tarefa. Ao usar uploads de várias partes dessa forma, o committer melhora o desempenho da confirmação da tarefa em relação à versão 2 do FileOutputCommitter algoritmo padrão. Ao usar o EMRFS S3 Optimized Committer, há algumas diferenças fundamentais de comportamento em relação ao comportamento tradicional de multipart uploads a considerar:

- Os multipart uploads são sempre executados, independentemente do tamanho do arquivo. Isso é diferente do comportamento padrão do EMRFS, em que a propriedade `fs.s3n.multipart.uploads.split.size` controla o tamanho do arquivo no qual multipart uploads são acionados.
- Os multipart uploads são deixados incompletos por um período mais longo até que a tarefa seja confirmada ou cancelada. Isso é diferente do comportamento padrão do EMRFS no qual um multipart upload é concluído quando uma tarefa é concluída ao gravar um determinado arquivo.

Devido a essas diferenças, se uma JVM do executor do Spark apresenta falha ou é eliminada enquanto as tarefas estão executando e gravando dados no Amazon S3, é mais provável que os carregamentos multipart partes sejam abandonados. Por esse motivo, quando você usa o EMRFS S3 Optimized Committer, certifique-se de seguir as melhores práticas para gerenciar multipart uploads com falha. Para obter mais informações, consulte Práticas recomendadas para trabalhar com buckets do Amazon S3 no Guia de gerenciamento do Amazon EMR.

Considerações sobre ajuste de tarefas

O EMRFS S3 Optimized Committer consome uma pequena quantidade de memória para cada arquivo gravado por uma tentativa de tarefa até que a tarefa seja confirmada ou cancelada. Na maioria dos trabalhos, a quantidade de memória consumida é insignificante. Para trabalhos que têm tarefas de longa execução que gravam um grande número de arquivos, a memória que o committer consome pode ser perceptível e exigem ajustes na memória alocada para executores do Spark. É

possível ajustar a memória do executor usando a propriedade `spark.executor.memory`. Como diretriz, uma única tarefa gravando 100.000 arquivos normalmente exigiria 100 MB adicionais de memória. Para obter mais informações, consulte Propriedades da aplicação na documentação de configuração do Apache Spark.

Ative o confirmador otimizado para EMRFS S3 para o Amazon EMR 5.19.0

Se você estiver usando o Amazon EMR 5.19.0, poderá definir manualmente a propriedade `spark.sql.parquet.fs.optimized.committer.optimization-enabled` como `true` quando criar um cluster ou no Spark, se estiver usando o Amazon EMR.

Habilitar o confirmador otimizado para EMRFS S3 quando criar um cluster

Use a classificação de configuração `spark-defaults` para definir a propriedade `spark.sql.parquet.fs.optimized.committer.optimization-enabled` como `true`. Para obter mais informações, consulte Configurar aplicações.

Habilitar o confirmador otimizado para EMRFS S3 no Spark

Você pode definir `spark.sql.parquet.fs.optimized.committer.optimization-enabled` como `true` codificando-o em um `SparkConf`, transmitindo-o como um parâmetro `--conf` no shell Spark ou nas ferramentas `spark-submit` e `spark-sql` ou em `conf/spark-defaults.conf`. Para obter mais informações, consulte Configuração do Spark na documentação do Apache Spark.

O exemplo a seguir mostra como habilitar o committer ao executar um comando spark-sql.

```
spark-sql \
  --conf spark.sql.parquet.fs.optimized.committer.optimization-enabled=true \
  -e "INSERT OVERWRITE TABLE target_table SELECT * FROM source_table;"
```

Use o protocolo de confirmação otimizada para EMRFS S3

O protocolo de confirmação otimizado para S3 do EMRFS é uma FileCommitProtocolimplementação alternativa otimizada para gravar arquivos com a substituição de partição dinâmica do Spark no Amazon S3 ao usar o EMRFS. O protocolo melhora a performance da aplicação ao evitar operações de renomeação no Amazon S3 durante a fase de confirmação do trabalho de substituição de partição dinâmica do Spark.

Observe que Usar o confirmador otimizado para EMRFS S3 também melhora o a performance ao evitar operações de renomeação. No entanto, ele não funciona para casos de substituição de

partição dinâmica, embora as melhorias no protocolo de confirmação sejam direcionados apenas a casos de substituição de partição dinâmica.

O confirmador está disponível com as versões 5.30.0 e posteriores e 6.2.0 e posteriores do Amazon EMR e é habilitado por padrão. O Amazon EMR adicionou uma melhoria no paralelismo a partir da versão 5.31.0. O protocolo é usado para trabalhos do Spark que usam Spark ou conjuntos de DataFrames dados. Há circunstâncias em que o protocolo de confirmação não é usado. Para obter mais informações, consulte Requisitos do protocolo de confirmação otimizado para EMRFS S3.

Tópicos

- Requisitos do protocolo de confirmação otimizado para EMRFS S3
- O protocolo de confirmação otimizado para EMRFS S3 e carregamentos multipart
- Considerações sobre ajuste de tarefas

Requisitos do protocolo de confirmação otimizado para EMRFS S3

O protocolo de confirmação otimizado para EMRFS S3 é usado quando as seguintes condições são atendidas:

- Você executa trabalhos do Spark que usam o Spark ou conjuntos de dados para substituir tabelas particionadas. DataFrames
- Você executa trabalhos do Spark cujo modo de substituição de partição é `dynamic`.
- Carregamentos multipart estão habilitados no Amazon EMR. Esse é o padrão. Para obter mais informações, consulte O protocolo de confirmação otimizado para EMRFS S3 e carregamentos multipart.
- O cache do sistema de arquivos para o EMRFS está habilitado. Esse é o padrão. Verifique se a configuração `fs.s3.impl.disable.cache` está definida como `false`.
- O suporte integrado de fonte de dados do Spark é usado. O suporte integrado ao Parquet é usado nas seguintes circunstâncias:
  - Quando os trabalhos gravam em fontes de dados ou tabelas integradas.
  - Quando os trabalhos gravam em tabelas do Parquet do metastore do Hive. Isso acontece quando `spark.sql.hive.convertInsertingPartitionedTable` e `spark.sql.hive.convertMetastoreParquet` são definidos como verdadeiros. Essas são as configurações padrão.
  - Quando os trabalhos gravam na tabela do ORC do metastore do Hive. Isso acontece quando `spark.sql.hive.convertInsertingPartitionedTable` e

`spark.sql.hive.convertMetastoreOrc` são definidos como `true`. Essas são as configurações padrão.

- As operações de trabalhos do Spark que gravam em um local de partição padrão, por exemplo, `${table_location}/k1=v1/k2=v2/`, usam o confirmador. O protocolo não será usado se uma operação de trabalho gravar em um local de partição personalizado, por exemplo, se o local de uma partição personalizado for definido usando o comando `ALTER TABLE SQL`.
- Os valores a seguir para o Spark devem ser usados:
  - `spark.sql.sources.commitProtocolClass` deve ser definido como `org.apache.spark.sql.execution.datasources.SQLEmrOptimizedCommitProtocol`. Essa é a configuração padrão para as versões 5.30.0 e superiores e 6.2.0 e superiores do Amazon EMR.
  - A opção de gravação `partitionOverwriteMode` ou `spark.sql.sources.partitionOverwriteMode` deve ser definida como `dynamic`. A configuração padrão é `static`.

    > **Note**
    >
    > A opção de gravação `partitionOverwriteMode` foi introduzida no Spark 2.4.0. Para a versão 2.3.2 do Spark, incluída com a versão 5.19.0 do Amazon EMR, defina a propriedade `spark.sql.sources.partitionOverwriteMode`.

  - Se os trabalhos do Spark substituírem a tabela do Parquet do metastore do Hive, `spark.sql.hive.convertMetastoreParquet`, `spark.sql.hive.convertInsertingPartitionedTable` e `spark.sql.hive.convertMetastore.partitionOverwriteMode` deverão ser configurados como `true`. Existem as configurações padrão.
  - Se os trabalhos do Spark substituírem a tabela do ORC do metastore do Hive, `spark.sql.hive.convertMetastoreOrc`, `spark.sql.hive.convertInsertingPartitionedTable` e `spark.sql.hive.convertMetastore.partitionOverwriteMode` deverão ser configurados como `true`. Existem as configurações padrão.

Example – Modo de substituição de partição dinâmica

Neste exemplo do Scala, a otimização é acionada. Primeiro, você define a propriedade `partitionOverwriteMode` como `dynamic`. Isso só substitui as partições nas quais você está

gravando dados. Em seguida, você especifica as colunas de partição dinâmica com `partitionBy` e define o modo de gravação como `overwrite`.

```
val dataset = spark.range(0, 10)
  .withColumn("dt", expr("date_sub(current_date(), id)"))

dataset.write.mode("overwrite")                 // "overwrite" instead of "insert"
  .option("partitionOverwriteMode", "dynamic")  // "dynamic" instead of "static"
  .partitionBy("dt")                            // partitioned data instead of
 unpartitioned data
  .parquet("s3://amzn-s3-demo-bucket1/output")    // "s3://" to use Amazon EMR file
 system, instead of "s3a://" or "hdfs://"
```

Quando o protocolo de confirmação otimizado para EMRFS S3 não é usado

Geralmente, o protocolo de confirmação otimizado para EMRFS S3 funciona da mesma forma que o protocolo de confirmação Spark padrão de código aberto,. `org.apache.spark.sql.execution.datasources.SQLHadoopMapReduceCommitProtocol` A otimização não ocorrerá nas situações a seguir.

Situação	Por que o protocolo de confirmação não é usado
Quando você grava no HDFS	O protocolo de confirmação só permite gravação no Amazon S3 com o uso do EMRFS.
Quando você usa o sistema de arquivos S3A	O protocolo de confirmação só é compatível com EMRFS.
Quando você usa MapReduce nossa API RDD do Spark	O protocolo de confirmação só oferece suporte ao uso de SparkSQL ou Dataset DataFrame. APIs
Quando a substituição da partição dinâmica não é acionada	O protocolo de confirmação só otimiza os casos de substituição de partição dinâmica. Para outros casos, consulte Usar o confirmador otimizado para EMRFS S3.

Os exemplos de Scala a seguir demonstram algumas situações adicionais que o protocolo de confirmação otimizado para EMRFS S3 delega para `SQLHadoopMapReduceCommitProtocol`.

Example – Modo de substituição de partição dinâmica com local de partição personalizado

Neste exemplo, os programas Scala sobrescrevem duas partições no modo de substituição dinâmica de partição. Uma partição tem um local de partição personalizado. A outra partição usa o local de partição padrão. O protocolo de confirmação otimizado para EMRFS S3 só melhora a partição que usa o local de partição padrão.

```
val table = "dataset"
val inputView = "tempView"
val location = "s3://bucket/table"

spark.sql(s"""
  CREATE TABLE $table (id bigint, dt date)
  USING PARQUET PARTITIONED BY (dt)
  LOCATION '$location'
""")

// Add a partition using a custom location
val customPartitionLocation = "s3://bucket/custom"
spark.sql(s"""
  ALTER TABLE $table ADD PARTITION (dt='2019-01-28')
  LOCATION '$customPartitionLocation'
""")

// Add another partition using default location
spark.sql(s"ALTER TABLE $table ADD PARTITION (dt='2019-01-29')")

def asDate(text: String) = lit(text).cast("date")

spark.range(0, 10)
  .withColumn("dt",
    when($"id" > 4, asDate("2019-01-28")).otherwise(asDate("2019-01-29")))
  .createTempView(inputView)

// Set partition overwrite mode to 'dynamic'
spark.sql(s"SET spark.sql.sources.partitionOverwriteMode=dynamic")

spark.sql(s"INSERT OVERWRITE TABLE $table SELECT * FROM $inputView")
```

O código Scala cria os seguintes objetos do Amazon S3:

```
custom/part-00001-035a2a9c-4a09-4917-8819-e77134342402.c000.snappy.parquet
custom_$folder$
table/_SUCCESS
table/dt=2019-01-29/part-00000-035a2a9c-4a09-4917-8819-e77134342402.c000.snappy.parquet
table/dt=2019-01-29_$folder$
table_$folder$
```

> **Note**
>
> Gravar em locais de partição personalizados em versões anteriores do Spark pode resultar em perda de dados. Neste exemplo, a partição `dt='2019-01-28'` seria perdida. Para obter mais detalhes, consulte SPARK-35106. Isso foi corrigido nas versões 5.33.0 e posteriores do Amazon EMR, excluindo 6.0.x e 6.1.x.

Ao gravar em partições em locais personalizados, o Spark usa um algoritmo de confirmação semelhante ao exemplo anterior, que é descrito abaixo. Como no exemplo anterior, o algoritmo resulta em renomeações sequenciais, o que pode afetar negativamente a performance.

O algoritmo no Spark 2.4.0 segue estas etapas:

1. Ao gravar a saída em uma partição em um local personalizado, as tarefas gravam em um arquivo no diretório de preparação do Spark, que é criado no local de saída final. O nome do arquivo inclui um UUID aleatório para evitar colisões de nomes de arquivos. A tentativa de tarefa controla cada arquivo junto com o caminho de saída final desejado.
2. Quando uma tarefa é concluída com êxito, ela fornece o driver com os arquivos e os caminhos desejados de saída final.
3. Depois que todas as tarefas forem concluídas, a fase de confirmação do trabalho renomeará sequencialmente todos os arquivos que foram gravados para partições em locais personalizados em seus caminhos de saída final.
4. O diretório de preparação é excluído antes de a fase de confirmação de trabalho ser concluída.

O protocolo de confirmação otimizado para EMRFS S3 e carregamentos multipart

Para usar a otimização para substituição de partições dinâmicas no protocolo de confirmação otimizado para EMRFS S3, os carregamentos multipart devem ser habilitados no Amazon EMR. Multipart uploads são habilitados por padrão. Você pode habilitá-los novamente, se necessário.

Para obter mais informações, consulte Configure multipart upload for Amazon S3 (Configurar o carregamento fracionado no Amazon S3) no Guia de gerenciamento do Amazon EMR.

Durante a substituição de partições dinâmicas, o protocolo de confirmação otimizado para EMRFS S3 usa as características semelhantes a transações de carregamentos multipart para garantir que os arquivos gravados por tentativas de tarefas apareçam apenas no local de saída do trabalho após a confirmação do trabalho. Ao usar carregamentos multipart dessa maneira, o protocolo de confirmação melhora a performance de confirmação de trabalhos em relação ao padrão `SQLHadoopMapReduceCommitProtocol`. Quando o protocolo de confirmação otimizado para EMRFS S3 é usado, há algumas diferenças fundamentais com relação ao comportamento tradicional de carregamentos multipart a serem consideradas:

- Os multipart uploads são sempre executados, independentemente do tamanho do arquivo. Isso é diferente do comportamento padrão do EMRFS, em que a propriedade `fs.s3n.multipart.uploads.split.size` controla o tamanho do arquivo no qual multipart uploads são acionados.
- Os multipart uploads são deixados incompletos por um período mais longo até que a tarefa seja confirmada ou cancelada. Isso é diferente do comportamento padrão do EMRFS no qual um multipart upload é concluído quando uma tarefa é concluída ao gravar um determinado arquivo.

Devido a essas diferenças, se uma JVM do executor do Spark apresenta falha ou é eliminada enquanto as tarefas estão executando e gravando dados no Amazon S3 ou se uma JVM do executor do Spark apresenta falha ou é eliminada enquanto um trabalho está sendo executado, é mais provável que os carregamentos multipart sejam abandonados. Por esse motivo, quando você usa o protocolo de confirmação otimizado para EMRFS S3, não deixe de seguir as práticas recomendadas para gerenciar carregamentos multipart que apresentaram falha. Para obter mais informações, consulte Práticas recomendadas para trabalhar com buckets do Amazon S3 no Guia de gerenciamento do Amazon EMR.

Considerações sobre ajuste de tarefas

Em executores do Spark, o protocolo de confirmação otimizado para EMRFS S3 consome uma pequena quantidade de memória para cada arquivo gravado por uma tentativa de tarefa até que a tarefa seja confirmada ou cancelada. Na maioria dos trabalhos, a quantidade de memória consumida é insignificante.

Nos drivers do Spark, o protocolo de confirmação otimizado para EMRFS S3 exige memória para armazenar informações de metadados de cada arquivo confirmado até que o trabalho seja

confirmado ou cancelado. Na maioria dos trabalhos, a configuração padrão de memória do driver do Spark é insignificante.

Para trabalhos que têm tarefas de execução prolongada que gravam um grande número de arquivos, a memória que o confirmador consome pode ser perceptível e precisar de ajustes na memória alocada para executores do Spark. Você pode ajustar a memória usando a `spark.driver.memory` propriedade dos drivers do Spark e a `spark.executor.memory` propriedade dos executores do Spark. Como diretriz, uma única tarefa gravando 100.000 arquivos normalmente exigiria 100 MB adicionais de memória. Para obter mais informações, consulte Propriedades da aplicação na documentação de configuração do Apache Spark.

Tentar novamente solicitações do Amazon S3 com EMRFS

Este tópico fornece informações sobre as estratégias de novas tentativas que você pode usar ao fazer solicitações ao Amazon S3 com o EMRFS. Quando a taxa de solicitação aumenta, o S3 tenta escalar para suportar a nova taxa. Durante esse processo, o S3 pode controlar a utilização das solicitações e retornar um erro `503 Slow Down`. Para melhorar a taxa de sucesso das solicitações do S3, você pode ajustar sua estratégia de novas tentativas configurando propriedades na sua configuração `emrfs-site`.

Você pode ajustar sua estratégia de novas tentativas das maneiras a seguir.

- Aumente o limite máximo de novas tentativas para a estratégia padrão de novas tentativas de recuo exponencial.
- Habilite e configure a estratégia de novas tentativas de additive-increase/multiplicative-decrease (AIMD - aumento aditivo/diminuição multiplicativa). O AIMD é compatível com versões 6.4.0 e posteriores do Amazon EMR.

Usar a estratégia padrão de recuo exponencial

Por padrão, o EMRFS usa uma estratégia de recuo exponencial para tentar novamente solicitações do Amazon S3. O limite padrão de novas tentativas do EMRFS é 15. Para evitar um erro `503 Slow Down` do S3, você pode aumentar o limite de novas tentativas ao criar um novo cluster, em um cluster em execução ou no runtime da aplicação.

Para aumentar o limite de novas tentativas, você deve alterar o valor de `fs.s3.maxRetries` na sua configuração `emrfs-site`. O exemplo de configuração a seguir define `fs.s3.maxRetries` como um valor personalizado de 30.

```
[
    {
      "Classification": "emrfs-site",
      "Properties": {
        "fs.s3.maxRetries": "30"
      }
    }
]
```

Para obter mais informações sobre como trabalhar com objetos de configuração, consulte Configurar aplicações.

Usar a estratégia de novas tentativas do AIMD

Com as versões 6.4.0 e posteriores do Amazon EMR, o EMRFS é compatível com uma estratégia alternativa de novas tentativas baseada em um modelo de aumento aditivo/diminuição multiplicativa (AIMD). A estratégia de novas tentativa do AIMD é especialmente útil quando você trabalha com grandes clusters do Amazon EMR.

O AIMD calcula uma taxa de solicitação personalizada usando dados sobre solicitações recentes bem-sucedidas. Essa estratégia diminui o número de solicitações submetidas a controle de utilização e o total de tentativas necessárias por solicitação.

Para habilitar a estratégia de novas tentativas do AIMD, você deve definir a propriedade `fs.s3.aimd.enabled` como `true` em sua configuração `emrfs-site`, como no exemplo a seguir.

```
[
    {
      "Classification": "emrfs-site",
      "Properties": {
        "fs.s3.aimd.enabled": "true"
      }
    }
]
```

Para obter mais informações sobre como trabalhar com objetos de configuração, consulte Configurar aplicações.

Configurações avançadas de novas tentativas do AIMD

Você pode configurar as propriedades listadas na tabela a seguir para refinar o comportamento de novas tentativas quando usar a estratégia de novas tentativas do AIMD. Para a maioria dos casos de uso, recomendamos que você use os valores padrão.

Propriedades avançadas da estratégia de novas tentativas do AIMD

Propriedade	Valor padrão	Descrição
`fs.s3.aimd.increaseIncrement`	0.1	Controla a rapidez com que a taxa de solicitações aumenta quando solicitações consecutivas são bem-sucedidas.
`fs.s3.aimd.reductionFactor`	2	Controla a rapidez com que a taxa de solicitação diminui quando o Amazon S3 retorna uma resposta 503. O fator padrão de 2 reduz a taxa de solicitação pela metade.
`fs.s3.aimd.minRate`	0.1	Define o limite inferior da taxa de solicitações quando as solicitações sofrem controle de utilização sustentado pelo S3.
`fs.s3.aimd.initialRate`	5500	Define a taxa de solicitação inicial, que sofre alterações de acordo com os valores que você especifica para `fs.s3.aimd.increaseIncrement` e `fs.s3.aimd.reductionFactor` . A taxa inicial também é usada para solicitações GET e é

Propriedade	Valor padrão	Descrição
		escalada proporcionalmente (3500/5500) para solicitações PUT.
`fs.s3.aimd.adjustWindow`	2	Controla a frequência com que a taxa de solicitação é ajustada, medida em número de respostas.
`fs.s3.aimd.maxAttempts`	100	Define o número máximo de tentativas para testar uma solicitação.

Adicionar uma etapa do Spark

Você pode usar etapas do Amazon EMR para enviar trabalhos à estrutura do Spark instalada em um cluster do EMR. Para obter mais informações, consulte Etapas no Guia de gerenciamento do Amazon EMR. No console e na CLI, você faz isso usando uma etapa de aplicativo Spark, que executa o script `spark-submit` como um etapa em seu nome. Com a API, você pode usar uma etapa para invocar `spark-submit` usando `command-runner.jar`.

Para obter mais informações sobre como enviar aplicações ao Spark, consulte o tópico Submitting applications na documentação do Apache Spark.

Para enviar uma etapa do Spark usando o console

1. Abra o console do Amazon EMR em https://console.aws.amazon.com /emr.
2. Em Cluster List (Lista de clusters), escolha o nome do cluster.
3. Role até a seção Steps (Etapas) e expanda-a. Em seguida, escolha Add step (Adicionar etapa).
4. Na caixa de diálogo Add Step (Adicionar etapa):
   - Para Step type (Tipo de etapa), escolha Spark application (Aplicativo Spark).
   - Para Name (Nome), aceite o nome padrão (aplicativo Spark) ou digite um novo nome.
   - Em Deploy mode (Modo de implantação), escolha o modo de Client (Cliente) ou de Cluster. O modo de cliente inicia o programa de driver na instância primária do cluster, enquanto o modo

de cluster inicia o programa de driver no cluster. Para o modo de cliente, a saída do log do driver aparece nos logs de etapa, enquanto para o modo de cluster, a saída do log do driver aparece nos logs do primeiro recipiente YARN. Para obter mais informações, consulte Cluster mode overview na documentação do Apache Spark.

- Especifique as Spark-submit options (opções de Spark-submit) desejadas. Para obter mais informações sobre as opções de `spark-submit`, consulte Launching applications with spark-submit.
- Para Application location (Local do aplicativo), especifique o local ou o caminho do URI do S3 do aplicativo.
- Para Arguments (Argumentos), deixe o campo em branco.
- Para Action on failure (Ação na falha), aceite a opção padrão Continue (Continuar).

5. Escolha Adicionar. A etapa é exibida no console com o status Pendente.
6. O status da etapa muda de Pending (Pendente) para Running (Em execução) e depois para Completed (Concluído) conforme ela é executada. Para atualizar o status, escolha o ícone Refresh (Atualizar) acima da coluna Actions (Ações).
7. Os resultados da etapa estão localizados na página Detalhes do cluster do console do Amazon EMR ao lado da etapa em Arquivos de log se você tem o registro em log configurado. Opcionalmente, você pode encontrar informações de etapas no bucket de logs configurado ao executar o cluster.

Para enviar trabalhos para o Spark usando o AWS CLI

Envie uma etapa ao criar o cluster ou use o subcomando `aws emr add-steps` em um cluster existente.

1. Use `create-cluster`, conforme mostrado no exemplo a seguir.

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr create-cluster --name "Add Spark Step Cluster" --release-label emr-7.9.0 --
applications Name=Spark \
```

```
--ec2-attributes KeyName=myKey --instance-type m5.xlarge --instance-count 3 \
--steps Type=Spark,Name="Spark Program",ActionOnFailure=CONTINUE,Args=[--
class,org.apache.spark.examples.SparkPi,/usr/lib/spark/examples/jars/spark-
examples.jar,10] --use-default-roles
```

Como opção, você pode usar `command-runner.jar` conforme mostrado no exemplo a seguir.

```
aws emr create-cluster --name "Add Spark Step Cluster" --release-label emr-7.9.0 \
--applications Name=Spark --ec2-attributes KeyName=myKey --instance-type m5.xlarge
 --instance-count 3 \
--steps Type=CUSTOM_JAR,Name="Spark Program",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[spark-example,SparkPi,10] --use-default-
roles
```

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

2. Como alternativa, adicione etapas a um cluster já em execução. Use `add-steps`.

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF --steps
 Type=Spark,Name="Spark Program",ActionOnFailure=CONTINUE,Args=[--
class,org.apache.spark.examples.SparkPi,/usr/lib/spark/examples/jars/spark-
examples.jar,10]
```

Como opção, você pode usar `command-runner.jar` conforme mostrado no exemplo a seguir.

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF --steps Type=CUSTOM_JAR,Name="Spark
 Program",Jar="command-runner.jar",ActionOnFailure=CONTINUE,Args=[spark-
example,SparkPi,10]
```

Enviar trabalhos ao Spark usando o SDK para Java

1. O exemplo a seguir mostra como adicionar uma etapa a um cluster com o Spark usando Java.

```
AWSCredentials credentials = new BasicAWSCredentials(accessKey, secretKey);
AmazonElasticMapReduce emr = new AmazonElasticMapReduceClient(credentials);
```

```
StepFactory stepFactory = new StepFactory();
AmazonElasticMapReduceClient emr = new AmazonElasticMapReduceClient(credentials);
AddJobFlowStepsRequest req = new AddJobFlowStepsRequest();
req.withJobFlowId("j-1K48XXXXXXHCB");

List<StepConfig> stepConfigs = new ArrayList<StepConfig>();

HadoopJarStepConfig sparkStepConf = new HadoopJarStepConfig()
   .withJar("command-runner.jar")
   .withArgs("spark-submit","--executor-memory","1g","--
class","org.apache.spark.examples.SparkPi","/usr/lib/spark/examples/jars/spark-
examples.jar","10");

StepConfig sparkStep = new StepConfig()
   .withName("Spark Step")
   .withActionOnFailure("CONTINUE")
   .withHadoopJarStep(sparkStepConf);

stepConfigs.add(sparkStep);
req.withSteps(stepConfigs);
AddJobFlowStepsResult result = emr.addJobFlowSteps(req);
```

2. Visualize os resultados da etapa examinando os logs da etapa. Você pode fazer isso no AWS Management Console caso de ter ativado o registro, escolhendo Etapas, selecionando sua etapa e, em Arquivos de log, escolhendo `stdout` ou`stderr`. Para ver os logs disponíveis, escolha View Logs (Exibir logs).

Substituir as definições de configuração padrão do Spark

Você pode querer substituir valores de configuração padrão do Spark para cada aplicativo. Você pode fazer isso ao enviar aplicativos usando uma etapa, que transmite essencialmente opções para `spark-submit`. Por exemplo, você pode querer alterar a memória alocada para um processo de executor modificando `spark.executor.memory`. Você poderia fornecer à opção `--executor-memory` um argumento semelhante ao seguinte:

```
spark-submit --executor-memory 1g --class org.apache.spark.examples.SparkPi /usr/lib/
spark/examples/jars/spark-examples.jar 10
```

Da mesma forma, você pode ajustar `--executor-cores` e `--driver-memory`. Em uma etapa, você forneceria os seguintes argumentos para a etapa:

```
--executor-memory 1g --class org.apache.spark.examples.SparkPi /usr/lib/spark/examples/
jars/spark-examples.jar 10
```

Você também pode ajustar as configurações que não possam não ter uma opção interna usando `--conf`. Para obter mais informações sobre outras configurações que são ajustáveis, consulte o tópico Dynamically loading Spark properties na documentação do Apache Spark.

Visualizar o histórico de aplicações do Spark

Você pode visualizar detalhes da interface de usuário do Spark, da aplicação YARN e do Tez usando a guia Interfaces de usuário da aplicação da página de detalhes de um cluster no console. As interfaces de usuário (IUs) das aplicações do Amazon EMR facilitam a solução de problemas e a análise de trabalhos ativos e do histórico de trabalhos.

Para obter mais informações, consulte Visualizar histórico de aplicações no Guia de gerenciamento do Amazon EMR.

Acesse a web do Spark UIs

Você pode visualizar a web do Spark UIs seguindo os procedimentos para criar um túnel SSH ou criar um proxy na seção Connect to the cluster no Amazon EMR Management Guide e, em seguida, navegando até o YARN do seu cluster. ResourceManager Escolha o link em Tracking UI (Interface do usuário de rastreamento) para o seu aplicativo. Se seu aplicativo estiver em execução, você verá ApplicationMaster. Isso o levará à interface do usuário web do aplicativo principal na porta 20888 onde o driver está localizado. O driver pode estar localizado no nó primário do cluster se você executa no modo de cliente YARN. Se você estiver executando um aplicativo no modo de cluster do YARN, o driver está localizado no do aplicativo no cluster. ApplicationMaster Se seu aplicativo estiver concluído, você verá o Histórico, que o levará ao número da porta da HistoryServer interface do usuário do Spark em 18080 do nó primário do cluster EMR. Isso é para aplicativos que já foram concluídos. Você também pode navegar até a HistoryServer interface do usuário do Spark diretamente em http: //:18080/*`master-public-dns-name`*.

Com as versões 5.25.0 ou posteriores do Amazon EMR, você pode acessar a interface de usuário do servidor de histórico do Spark no console sem configurar um proxy da Web por meio de uma conexão SSH. Para obter mais informações, consulte View Persistent Application User Interfaces.

Uso do conector do Amazon Kinesis Data Streams para streaming estruturado do Spark

As versões 7.1.0 e posteriores do Amazon EMR incluem um conector do Amazon Kinesis Data Streams para streaming estruturado do Spark na imagem de lançamento. Com esse conector, você pode usar o Spark no Amazon EMR para processar dados armazenados no Amazon Kinesis Data Streams. O conector é compatível com os tipos de consumidor de `GetRecords` (throughput compartilhado) e `SubscribeToShard` (fan-out aprimorado). Essa integração é baseada no `spark-sql-kinesis-connector`. Para obter detalhes sobre como começar a usar o conector, consulte o README.

O exemplo a seguir demonstra como usar o conector para iniciar uma aplicação do Spark com o Amazon EMR

```
spark-submit my_kinesis_streaming_script.py
```

Usar a integração do Amazon Redshift para Apache Spark com o Amazon EMR

Com as versões 6.4.0 e posteriores do Amazon EMR, cada imagem de versão inclui um conector entre o Apache Spark e o Amazon Redshift. Com esse conector, você pode usar o Spark no Amazon EMR para processar dados armazenados no Amazon Redshift. Para as versões 6.4.0 a 6.8.0 do Amazon EMR, a integração é baseada no conector de código aberto `spark-redshift`. Para as versões 6.9.0 e posteriores do Amazon EMR, a integração do Amazon Redshift para Apache Spark foi migrada da versão comunitária para uma integração nativa.

Tópicos

- Inicialização de uma aplicação do Spark usando a integração do Amazon Redshift para Apache Spark
- Autenticação com a integração do Amazon Redshift para Apache Spark
- Leitura e gravação de e para o Amazon Redshift
- Considerações e limitações ao usar o conector do Spark

Inicialização de uma aplicação do Spark usando a integração do Amazon Redshift para Apache Spark

Para as versões 6.4 a 6.9 do Amazon EMR, você deve usar a opção `--jars` ou `--packages` para especificar quais dos seguintes arquivos JAR você deseja usar. A opção `--jars` especifica dependências armazenadas localmente, no HDFS ou usando HTTP/S. Para ver outros locais de arquivos compatíveis com a opção `--jars`, consulte Gerenciamento avançado de dependências na documentação do Spark. A opção `--packages` especifica dependências armazenadas no repositório público Maven.

- `spark-redshift.jar`
- `spark-avro.jar`
- `RedshiftJDBC.jar`
- `minimal-json.jar`

As versões 6.10.0 e superiores do Amazon EMR não exigem a dependência `minimal-json.jar` e, por padrão, instalam automaticamente as outras dependências em cada cluster. Os exemplos a seguir mostram como iniciar uma aplicação do Spark com a integração do Amazon Redshift para Apache Spark.

Amazon EMR 6.10.0 +

O exemplo a seguir mostra como iniciar uma aplicação do Spark com o conector `spark-redshift` com as versões 6.10 e superiores do Amazon EMR.

```
spark-submit my_script.py
```

Amazon EMR 6.4.0 - 6.9.x

Para iniciar uma aplicação do Spark com o conector `spark-redshift` nas versões 6.4 a 6.9 do Amazon EMR, você deve usar a opção `--jars` ou `--packages`, conforme mostrado no exemplo a seguir. Observe que os caminhos listados com a opção `--jars` são os caminhos padrão para os arquivos JAR.

```
spark-submit \
  --jars /usr/share/aws/redshift/jdbc/RedshiftJDBC.jar,/usr/share/aws/redshift/
spark-redshift/lib/spark-redshift.jar,/usr/share/aws/redshift/spark-redshift/lib/
spark-avro.jar,/usr/share/aws/redshift/spark-redshift/lib/minimal-json.jar \
```

```
  my_script.py
```

Autenticação com a integração do Amazon Redshift para Apache Spark

Usando AWS Secrets Manager para recuperar credenciais e conectar-se ao Amazon Redshift

O exemplo de código a seguir mostra como você pode usar AWS Secrets Manager para recuperar credenciais para se conectar a um cluster do Amazon Redshift com a interface PySpark do Apache Spark em Python.

```
from pyspark.sql import SQLContext
import boto3

sc = # existing SparkContext
sql_context = SQLContext(sc)

secretsmanager_client = boto3.client('secretsmanager')
secret_manager_response = secretsmanager_client.get_secret_value(
    SecretId='string',
    VersionId='string',
    VersionStage='string'
)
username = # get username from secret_manager_response
password = # get password from secret_manager_response
url = "jdbc:redshift://redshifthost:5439/database?user=" + username + "&password=" +
 password

# Read data from a table
df = sql_context.read \
    .format("io.github.spark_redshift_community.spark.redshift") \
    .option("url", url) \
    .option("dbtable", "my_table") \
    .option("tempdir", "s3://path/for/temp/data") \
    .load()
```

Usar o IAM para recuperar credenciais e conectar-se ao Amazon Redshift

Você pode usar a versão 2 do driver JDBC fornecido pelo Amazon Redshift para se conectar ao Amazon Redshift com o conector do Spark. Para usar AWS Identity and Access Management

(IAM), configure seu URL do JDBC para usar a autenticação do IAM. Para se conectar a um cluster do Redshift pelo Amazon EMR, você deve dar ao perfil do IAM as permissões para recuperar credenciais temporárias do IAM. Atribua as permissões ao perfil do IAM para que ele possa recuperar credenciais e executar operações do Amazon S3.

- Redshift: GetClusterCredentials (para clusters provisionados do Amazon Redshift)
- Redshift: DescribeClusters (para clusters provisionados do Amazon Redshift)
- Redshift: GetWorkgroup (para grupos de trabalho sem servidor do Amazon Redshift)
- Redshift: GetCredentials (para grupos de trabalho sem servidor do Amazon Redshift)
- s3: GetBucket
- s3: GetBucketLocation
- s3: GetObject
- s3: PutObject
- s3: GetBucketLifecycleConfiguration

Para obter mais informações sobre `GetClusterCredentials`, consulte Políticas de recursos para `GetClusterCredentials`.

Você também deve garantir que o Amazon Redshift possa assumir o perfil do IAM durante as operações `COPY` e `UNLOAD`.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "Service": "redshift.amazonaws.com"
            },
            "Action": "sts:AssumeRole"
        }
    ]
}
```

O exemplo a seguir usa a autenticação do IAM entre o Spark e o Amazon Redshift:

```
from pyspark.sql import SQLContext
import boto3
```

```
sc = # existing SparkContext
sql_context = SQLContext(sc)

url = "jdbc:redshift:iam://redshift-host:redshift-port/db-name"
iam_role_arn = "arn:aws:iam::account-id:role/role-name"

# Read data from a table
df = sql_context.read \
    .format("io.github.spark_redshift_community.spark.redshift") \
    .option("url", url) \
    .option("aws_iam_role", iam_role_arn) \
    .option("dbtable", "my_table") \
    .option("tempdir", "s3a://path/for/temp/data") \
    .mode("error") \
    .load()
```

Leitura e gravação de e para o Amazon Redshift

Os exemplos de código a seguir são usados PySpark para ler e gravar dados de amostra de e para um banco de dados do Amazon Redshift com API de fonte de dados e usando o SparkSQL.

Data source API

Use PySpark para ler e gravar dados de amostra de e para um banco de dados do Amazon Redshift com a API de fonte de dados.

```
import boto3
from pyspark.sql import SQLContext

sc = # existing SparkContext
sql_context = SQLContext(sc)

url = "jdbc:redshift:iam://redshifthost:5439/database"
aws_iam_role_arn = "arn:aws:iam::accountID:role/roleName"

df = sql_context.read \
    .format("io.github.spark_redshift_community.spark.redshift") \
    .option("url", url) \
    .option("dbtable", "tableName") \
    .option("tempdir", "s3://path/for/temp/data") \
    .option("aws_iam_role", "aws_iam_role_arn") \
    .load()
```

```
df.write \
    .format("io.github.spark_redshift_community.spark.redshift") \
    .option("url", url) \
    .option("dbtable", "tableName_copy") \
    .option("tempdir", "s3://path/for/temp/data") \
    .option("aws_iam_role", "aws_iam_role_arn") \
    .mode("error") \
    .save()
```

SparkSQL

Use PySpark para ler e gravar dados de amostra de e para um banco de dados do Amazon Redshift com o SparkSQL.

```
import boto3
import json
import sys
import os
from pyspark.sql import SparkSession

spark = SparkSession \
    .builder \
    .enableHiveSupport() \
    .getOrCreate()

url = "jdbc:redshift:iam://redshifthost:5439/database"
aws_iam_role_arn = "arn:aws:iam::accountID:role/roleName"

bucket = "s3://path/for/temp/data"
tableName = "tableName" # Redshift table name

s = f"""CREATE TABLE IF NOT EXISTS {tableName} (country string, data string)
    USING io.github.spark_redshift_community.spark.redshift
    OPTIONS (dbtable '{tableName}', tempdir '{bucket}', url '{url}', aws_iam_role
 '{aws_iam_role_arn}' ); """

spark.sql(s)

columns = ["country" ,"data"]
data = [("test-country","test-data")]
df = spark.sparkContext.parallelize(data).toDF(columns)
```

```
# Insert data into table
df.write.insertInto(tableName, overwrite=False)
df = spark.sql(f"SELECT * FROM {tableName}")
df.show()
```

Considerações e limitações ao usar o conector do Spark

- Recomendamos que você ative o SSL para a conexão JDBC do Spark no Amazon EMR ao Amazon Redshift.
- Recomendamos que você gerencie as credenciais do cluster do Amazon Redshift no AWS Secrets Manager como uma prática recomendada. Consulte Usando AWS Secrets Manager para recuperar credenciais para se conectar ao Amazon Redshift para ver um exemplo.
- Recomendamos que você transmita um perfil do IAM com o parâmetro `aws_iam_role` para o parâmetro de autenticação do Amazon Redshift.
- O URI `tempdir` aponta para um local do Amazon S3. Esse diretório temporário não é limpo automaticamente e, portanto, pode incorrer em custos adicionais.
- Considere as seguintes recomendações para o Amazon Redshift:
  - Recomendamos bloquear o acesso público ao cluster do Amazon Redshift.
  - Recomendamos ativar o registro em log de auditoria do Amazon Redshift.
  - Recomendamos que você ative a Criptografia em repouso do Amazon Redshift.
- Considere as seguintes recomendações para o Amazon S3:
  - Recomendamos que você bloqueie o acesso público aos buckets do Amazon S3.
  - Recomendamos que você use criptografia no lado do servidor do Amazon S3 para criptografar os buckets do Amazon S3 usados.
  - Recomendamos que você use as políticas de ciclo de vida do Amazon S3 para definir as regras de retenção para o bucket do Amazon S3.
  - O Amazon EMR sempre verifica o código importado do código aberto para a imagem. Por motivos de segurança, não oferecemos suporte aos seguintes métodos de autenticação do Spark para o Amazon S3:
    - Definindo chaves de AWS acesso na classificação `hadoop-env` de configuração
    - Codificando chaves de AWS acesso no URI `tempdir`

Para obter mais informações sobre como usar o conector e os parâmetros compatíveis, consulte os seguintes recursos:

- Integração do Amazon Redshift para Apache Spark no Guia de gerenciamento do Amazon Redshift.
- O repositório da comunidade `spark-redshift` no GitHub.

Histórico de versões do Spark

A tabela a seguir informa a versão do Spark incluída em cada versão do Amazon EMR e os componentes instalados com a aplicação. Para obter as versões dos componentes em cada versão, consulte a seção Versão de componentes da sua versão em Versões 7.x do Amazon EMR, Versões de lançamento 6.x do Amazon EMR ou Versões de lançamento 5.x do Amazon EMR.

> **Important**
>
> A versão 2.3.1 do Apache Spark, disponível desde a versão 5.16.0 do Amazon EMR, aborda CVE-2018-8024 e CVE-2018-1334. Recomendamos que você migre as versões anteriores do Spark para a versão 2.3.1 ou posteriores.

Informações das versões do Spark

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-7.9.0	3.5.5	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
		history-server, spark-on-yarn, spark-yarn-slave
emr-7.8.0	3.5.4	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-7.7.0	3.5.3	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-7.6.0	3.5.3	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-7.5.0	3.5.2	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-7.4.0	3.5.2	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-7.3.0	3.5.1	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-7.2.0	3.5.1	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.36.2	2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-7.1.0	3.5.0	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-7.0.0	3.5.0	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-6.15.0	3.4.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.14.0	3.4.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-6.13.0	3.4.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.12.0	3.4.0	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-6.11.1	3.3.2	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.11.0	3.3.2	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-6.10.1	3.3.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.10.0	3.3.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-6.9.1	3.3.0	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.9.0	3.3.0	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-6.8.1	3.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.8.0	3.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-6.7.0	3.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.36.1	2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.36.0	2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.6.0	3.2.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.35.0	2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.5.0	3.1.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-6.4.0	3.1.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.3.1	3.1.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-6.3.0	3.1.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.2.1	3.0.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-6.2.0	3.0.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.1.1	3.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-6.1.0	3.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.0.1	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-6.0.0	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.34.0	2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.33.1	2.4.7	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.33.0	2.4.7	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.32.1	2.4.7	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.32.0	2.4.7	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.31.1	2.4.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.31.0	2.4.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.30.2	2.4.5	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.30.1	2.4.5	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.30.0	2.4.5	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-notebook-env, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.29.0	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.28.1	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.28.0	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.27.1	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.27.0	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.26.0	2.4.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.25.0	2.4.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.24.1	2.4.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.24.0	2.4.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.23.1	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.23.0	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.22.0	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.21.2	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.21.1	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.21.0	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.20.1	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.20.0	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.19.1	2.3.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.19.0	2.3.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.18.1	2.3.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.18.0	2.3.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.17.2	2.3.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.17.1	2.3.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.17.0	2.3.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.16.1	2.3.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.16.0	2.3.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.15.1	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.15.0	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.14.2	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.14.1	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.14.0	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.13.1	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.13.0	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.12.3	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.12.2	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.12.1	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.12.0	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.11.4	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.11.3	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.11.2	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.11.1	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.11.0	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.10.1	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.10.0	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.9.1	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.9.0	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.8.3	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.8.2	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.8.1	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.8.0	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.7.1	2.1.1	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.7.0	2.1.1	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.6.1	2.1.1	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.6.0	2.1.1	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.5.4	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.5.3	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.5.2	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.5.1	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.5.0	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.4.1	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.4.0	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.3.2	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.3.1	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.3.0	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.2.3	2.0.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.2.2	2.0.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.2.1	2.0.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.2.0	2.0.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.1.1	2.0.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.1.0	2.0.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.0.3	2.0.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-5.0.2	2.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.0.1	2.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.0.0	2.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-4.9.6	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.9.5	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.9.4	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-4.9.3	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.9.2	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.9.1	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-4.8.5	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.8.4	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.8.3	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-4.8.2	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.8.1	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.8.0	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-4.7.4	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.7.3	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.7.2	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-4.7.1	1.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.7.0	1.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.6.1	1.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-4.6.0	1.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.5.0	1.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.4.0	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-4.3.0	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.2.0	1.5.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.1.0	1.5.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rótulo de versão do Amazon EMR	Versão do Spark	Componentes instalados com o Spark
emr-4.0.0	1.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Apache Sqoop

> **Note**
>
> A última versão a incluir o Apache Sqoop foi a versão 7.4 do Amazon EMR. O Apache Sqoop será removido do Amazon EMR versão 7.5 e versões superiores.

O Apache Sqoop é uma ferramenta para transferência de dados entre os bancos de dados do Amazon S3, do Hadoop, do HDFS e do RDBMS. Para obter mais informações, consulte o site do Apache Sqoop. O Sqoop está incluído nas versões 5.0.0 e posteriores do Amazon EMR. As versões anteriores incluem o Sqoop como uma aplicação em sandbox. Para obter mais informações, consulte Versões de lançamento 4.x do Amazon EMR.

Tópicos

- Informações de versão do Sqoop
- Considerações sobre o Sqoop no Amazon EMR
- Histórico de versões do Sqoop

Informações de versão do Sqoop

Versão Sqoop para a versão 7.4.

A tabela a seguir lista a versão do Sqoop incluída na versão mais recente da série 7.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Sqoop.

Informações sobre a versão do Spark para o emr-7.4.0

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-7.4.0	Sqoop 1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
		server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Versão do Sqoop para 6.15.0

A tabela a seguir lista a versão do Sqoop incluída na versão mais recente da série 6.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Sqoop.

Para obter a versão dos componentes instalados com o Sqoop nessa versão, consulte Release 6.15.0 Component Versions.

Informações de versão do Sqoop para o emr-6.15.0

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-6.15.0	Sqoop 1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Versão do Sqoop para 5.36.2

A tabela a seguir lista a versão do Sqoop incluída na versão mais recente da série 5.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Sqoop.

Para obter a versão dos componentes instalados com o Sqoop nesse lançamento, consulte Release 5.36.2 Component Versions.

Informações de versão do Sqoop para o emr-5.36.2

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.36.2	Sqoop 1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Considerações sobre o Sqoop no Amazon EMR

Considere os seguintes itens quando executar o Sqoop no Amazon EMR.

Usando o Sqoop com integração HCatalog

O Sqoop no Amazon EMR oferece suporte à integração com o Sqoop. HCatalog
Ao usar o Sqoop para gravar a saída em uma HCatalog tabela no Amazon S3, desative a gravação direta do Amazon EMR definindo as propriedades e como. `mapred.output.direct.NativeS3FileSystem` `mapred.output.direct.EmrFileSystem` `false` Para obter mais informações, consulte Usando HCatalog. Você pode usar os comandos `-D mapred.output.direct.NativeS3FileSystem=false` e `-D mapred.output.direct.EmrFileSystem=false` do Hadoop. Se você não desabilitar a gravação direta, nenhum erro ocorrerá, mas a tabela será criada no Amazon S3 e nenhum dado será gravado.

Suporte do JDBC e de banco de dados para o Sqoop

Por padrão, o Sqoop tem um driver MariaDB e PostgreSQL instalado. O driver do PostgreSQL instalado para o Sqoop só funciona para o PostgreSQL 8.4. Para instalar um conjunto alternativo de conectores JDBC para o Sqoop, conecte-se ao nó principal do cluster e instale-o em `/usr/lib/sqoop/lib`. Veja a seguir links para vários conectores JDBC:

- MariaDB: Sobre o MariaDB Connector/J.
- PostgreSQL: Driver JDBC para PostgreSQL.
- SQLServer: Baixe o driver Microsoft JDBC para SQL Server.
- MySQL: baixar o Connector/J
- Oracle: obter drivers Oracle JDBC e UCP no repositório Maven do Oracle

Os bancos de dados compatíveis com o Sqoop estão listados no seguinte url, http://sqoop.apache.org/docs/*version*/SqoopUserGuide.html #_supported_databases, onde *version* está a versão do Sqoop que você está usando, por exemplo, 1.4.6. Se a string de conexão JDBC não corresponder às desta lista, você precisará especificar um driver.

Por exemplo, é possível exportar para uma tabela de banco de dados do Amazon Redshift com o seguinte comando (para JDBC 4.1):

```
sqoop export --connect jdbc:redshift://$MYREDSHIFTHOST:5439/mydb --
table mysqoopexport --export-dir s3://amzn-s3-demo-bucket/myinputfiles/ --driver
 com.amazon.redshift.jdbc41.Driver --username master --password Mymasterpass1
```

Você pode usar ambas as strings de conexão MariaDB e MySQL, mas, se especificar a string de conexão MariaDB, precisará especificar o driver:

```
sqoop export --connect jdbc:mariadb://$HOSTNAME:3306/mydb --table mysqoopexport --
export-dir s3://amzn-s3-demo-bucket/myinputfiles/ --driver org.mariadb.jdbc.Driver --
username master --password Mymasterpass1
```

Se você estiver usando a criptografia SSL para acessar seu banco de dados, precisará usar um URI do JDBC, como no seguinte exemplo de exportação do Sqoop:

```
sqoop export --connect jdbc:mariadb://$HOSTNAME:3306/mydb?
verifyServerCertificate=false&useSSL=true&requireSSL=true --table mysqoopexport --
```

```
export-dir s3://amzn-s3-demo-bucket/myinputfiles/ --driver org.mariadb.jdbc.Driver --
username master --password Mymasterpass1
```

Para obter mais informações sobre criptografia SSL no RDS, consulte Usar o SSL para criptografar uma conexão com uma instância de banco de dados no Guia do usuário do Amazon RDS.

Para obter mais informações, consulte a documentação do Apache Sqoop.

Proteção da senha

Há vários métodos que você pode escolher para transmitir sua senha com segurança:

Java KeyStore

O método preferido criptografa a senha com um Java KeyStore (JKS), eliminando a necessidade de armazenar a senha em um formato legível.

1. Crie um alias de senha. Quando solicitado, insira a senha usada para acessar o banco de dados.

   ```
   hadoop credential create mydb.password.alias -provider jceks://hdfs/user/root/
   mysql.password.jceks
   ```

2. Use o alias de senha para executar o trabalho do Sqoop:

   ```
   sqoop export -Dhadoop.security.credential.provider.path=jceks://hdfs/user/
   root/mysql.password.jceks --connect jdbc:mariadb://$HOSTNAME:3306/mydb --table
    mysqoopexport --export-dir s3://amzn-s3-demo-bucket/myinputfiles/ --driver
    org.mariadb.jdbc.Driver --username master --password-alias mydb.password.alias
   ```

--password-file

É possível usar o comando `--password-file` para transmitir a senha por meio de um arquivo, conforme mostrado no seguinte exemplo:

1. Crie um arquivo que contenha a senha:

   ```
   echo -n 'Mymasterpass1' > /home/hadoop/mysql-pass.password
   ```

2. Use o arquivo para executar o trabalho do Sqoop:

```
sqoop export --connect jdbc:mariadb://$HOSTNAME:3306/mydb --table
 mysqoopexport --export-dir s3://amzn-s3-demo-bucket/myinputfiles/ --driver
 org.mariadb.jdbc.Driver --username master --password-file /home/hadoop/mysql-
pass.password
```

-P

É possível usar o comando `-P` para inserir a senha por meio de um prompt, conforme mostrado no seguinte exemplo:

```
sqoop export --connect jdbc:mariadb://$HOSTNAME:3306/mydb --table mysqoopexport --
export-dir s3://amzn-s3-demo-bucket/myinputfiles/ --driver org.mariadb.jdbc.Driver
 --username master -P
```

Histórico de versões do Sqoop

A tabela a seguir informa a versão do Sqoop incluída em cada versão do Amazon EMR e os componentes instalados com a aplicação. Para obter as versões dos componentes em cada versão, consulte a seção Versão de componentes da sua versão em Versões 7.x do Amazon EMR, Versões de lançamento 6.x do Amazon EMR ou Versões de lançamento 5.x do Amazon EMR.

Informações de versão do Sqoop

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-7.4.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
		timeline-server, mariadb-server, sqoop-client
emr-7.3.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-7.2.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.36.2	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-7.1.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-7.0.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.15.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-6.14.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.13.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-6.12.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.11.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-6.11.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.10.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-6.10.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.9.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-6.9.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.8.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-6.8.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.7.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.36.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.36.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-6.6.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.35.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-6.5.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.4.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-6.3.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.3.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-6.2.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.2.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-6.1.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.1.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.34.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.33.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.33.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.32.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.32.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.31.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.31.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.30.2	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.30.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.30.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.29.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.28.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.28.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.27.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.27.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.26.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.25.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.24.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.24.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.23.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.23.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.22.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.21.2	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.21.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.21.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.20.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.20.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.19.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.19.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.18.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.18.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.17.2	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.17.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.17.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.16.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.16.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.15.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.15.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.14.2	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.14.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.14.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.13.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.13.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.12.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.12.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.12.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.12.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.11.4	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.11.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.11.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.11.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.11.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.10.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.10.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.9.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.9.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.8.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.8.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.8.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.8.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.7.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.7.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.6.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.6.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.5.4	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.5.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.5.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.5.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.5.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.4.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.4.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.3.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.3.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.3.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.2.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.2.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.2.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.2.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.1.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.1.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.0.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.0.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.0.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Rótulo de versão do Amazon EMR	Versão do Sqoop	Componentes instalados com o Sqoop
emr-5.0.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

TensorFlow

TensorFlow é uma biblioteca de matemática simbólica de código aberto para aplicações de inteligência de máquina e aprendizado profundo. Para obter mais informações, consulte o TensorFlow site. TensorFlow está disponível com as versões 5.17.0 e posteriores do Amazon EMR.

A tabela a seguir lista a versão TensorFlow incluída na versão mais recente da série 7.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Amazon EMR. TensorFlow

Para obter a versão dos componentes instalados com TensorFlow essa versão, consulte Release 7.0 Component Versions.

TensorFlow informações de versão do emr-7.0

Rótulo de versão do Amazon EMR	TensorFlow Versão	Componentes instalados com TensorFlow
emr-7.0	TensorFlow 2.16.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

A tabela a seguir lista a versão TensorFlow incluída na versão mais recente da série 6.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Amazon EMR. TensorFlow

Para obter a versão dos componentes instalados com TensorFlow essa versão, consulte Release 6.15.0 Component Versions.

TensorFlow informações de versão do emr-6.0

Rótulo de versão do Amazon EMR	TensorFlow Versão	Componentes instalados com TensorFlow
emr-6.15.0	TensorFlow 2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

A tabela a seguir lista a versão TensorFlow incluída na versão mais recente da série 5.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Amazon EMR. TensorFlow

Para obter a versão dos componentes instalados com TensorFlow essa versão, consulte Release 5.36.2 Component Versions.

TensorFlow informações de versão do emr-5.2

Rótulo de versão do Amazon EMR	TensorFlow Versão	Componentes instalados com TensorFlow
emr-5.36.2	TensorFlow 2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

TensorFlow compilações por tipo de EC2 instância da Amazon

O Amazon EMR usa compilações diferentes da TensorFlow biblioteca de acordo com os tipos de instância que você escolhe para o cluster. O Amazon EMR oferece suporte TensorFlow para clusters com tipos de instância aarch64 (Graviton) para EMR-7.5.0 e superior. A tabela a seguir lista as compilações por tipo de instância.

EC2 tipos de instância	TensorFlow construir
M5 e C5	Tensorflow 2.16.1 com otimização do Intel MKL
P2, P4, P5, P5, G5, G5, G5, G5, G5, G5, G5, G5, G5, G5 GR6	Tensorflow 2.16.1 com CUDA 12.3, cuDNN 8.9.7.29
P3, P3DN, G3 e G3S	Tensorflow 2.16.1 com CUDA 12.3, cuDNN 8.9.7.29, NCCL 2.20.3-1 O Nvidia NCCL só está disponível apenas em instâncias P3. Contrato de licença do usuário final (EULA): ao usar componentes Nvidia no Amazon EMR, você aceita os termos e condições descritos no EULA do produto.
Todos os outros, exceto instâncias do Graviton	Tensorflow 2.16.1

Segurança

Além de seguir as orientações em Usar TensorFlow com segurança, recomendamos que você inicie o cluster em uma sub-rede privada para ajudar você a limitar o acesso a fontes confiáveis. Para obter mais informações, consulte Amazon VPC options no Guia de gerenciamento do Amazon EMR.

Usando TensorBoard

TensorBoard é um conjunto de ferramentas de visualização para TensorFlow programas. Para obter mais informações, consulte TensorBoard: Aprendizado visualizado no site do Tensorflow.

Para usar TensorBoard com o Amazon EMR, você deve começar TensorBoard no nó principal do cluster.

Usar o Tensorboard com o Tensorflow no Amazon EMR

1. Conecte-se ao nó principal do cluster usando SSH. Para obter mais informações, consulte Conectar-se ao nó principal usando SSH no Guia de gerenciamento do Amazon EMR.
2. Digite o seguinte comando para iniciar o Tensorboard no nó principal. Substitua */my/log/directory* por um diretório no nó principal o qual você tiver gerado e armazenado os dados de resumo usando um gravador de resumo.

   Amazon EMR 5.19.0 and later

   ```
   python3 -m tensorboard.main --logdir=/home/hadoop/tensor --bind_all
   ```

   Amazon EMR 5.18.1 and earlier

   ```
   python3 -m tensorboard.main --logdir=/my/log/dir
   ```

   Por padrão, o nó principal hospeda TensorBoard usando a porta 6006 e o nome DNS público principal. Depois de iniciar TensorBoard, a saída da linha de comando apresenta a URL que pode ser usada para se conectar TensorBoard, conforme mostrado no exemplo a seguir:

   ```
   TensorBoard 2.16.1 at http://master-public-dns-name:6006 (Press CTRL+C to quit)
   ```
3. Configure o acesso a interfaces web no nó principal a partir de clientes confiáveis. Para obter mais informações, consulte Visualizar interfaces da Web hospedadas em clusters do Amazon EMR no Guia de gerenciamento do Amazon EMR.
4. Aberto TensorBoard em`http://master-public-dns-name:6006`.

TensorFlow histórico de versões

A tabela a seguir lista a versão do TensorFlow Amazon EMR, junto com os componentes instalados com a aplicação. Para obter as versões dos componentes em cada versão, consulte a seção Versão de componentes da sua versão em Versões 7.x do Amazon EMR, Versões de lançamento 6.x do Amazon EMR ou Versões de lançamento 5.x do Amazon EMR.

TensorFlow informações sobre a versão

Rótulo de versão do Amazon EMR	TensorFlow Versão	Componentes instalados com TensorFlow
emr-7.0	2.16.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.0	2.16.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.7.0	2.16.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.6.0	2.16.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-

Rótulo de versão do Amazon EMR	TensorFlow Versão	Componentes instalados com TensorFlow
		hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.5.0	2.16.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.4.0	2.16.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Rótulo de versão do Amazon EMR	TensorFlow Versão	Componentes instalados com TensorFlow
emr-7.3.0	2.16.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.2.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.36.2	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Rótulo de versão do Amazon EMR	TensorFlow Versão	Componentes instalados com TensorFlow
emr-7.1.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.0.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.15.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Rótulo de versão do Amazon EMR	TensorFlow Versão	Componentes instalados com TensorFlow
emr-6.14.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.13.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.12.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Rótulo de versão do Amazon EMR	TensorFlow Versão	Componentes instalados com TensorFlow
emr-6.11.1	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.11.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.10.1	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Rótulo de versão do Amazon EMR	TensorFlow Versão	Componentes instalados com TensorFlow
emr-6.10.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.9.1	2.10.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.9.0	2.10.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Rótulo de versão do Amazon EMR	TensorFlow Versão	Componentes instalados com TensorFlow
emr-6.8.1	2.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.8.0	2.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.7.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Rótulo de versão do Amazon EMR	TensorFlow Versão	Componentes instalados com TensorFlow
emr-5.36.1	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.36.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.6.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Rótulo de versão do Amazon EMR	TensorFlow Versão	Componentes instalados com TensorFlow
emr-5.35.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.5.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.4.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Rótulo de versão do Amazon EMR	TensorFlow Versão	Componentes instalados com TensorFlow
emr-6.3.1	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.3.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.2.1	2.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Rótulo de versão do Amazon EMR	TensorFlow Versão	Componentes instalados com TensorFlow
emr-6.2.0	2.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.1.1	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.1.0	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Rótulo de versão do Amazon EMR	TensorFlow Versão	Componentes instalados com TensorFlow
emr-6.0.1	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.0.0	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.34.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Rótulo de versão do Amazon EMR	TensorFlow Versão	Componentes instalados com TensorFlow
emr-5.33.1	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.33.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.32.1	2.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Rótulo de versão do Amazon EMR	TensorFlow Versão	Componentes instalados com TensorFlow
emr-5.32.0	2.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.31.1	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.31.0	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Rótulo de versão do Amazon EMR	TensorFlow Versão	Componentes instalados com TensorFlow
emr-5.30.2	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.30.1	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.30.0	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Rótulo de versão do Amazon EMR	TensorFlow Versão	Componentes instalados com TensorFlow
emr-5.29.0	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.28.1	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.28.0	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Rótulo de versão do Amazon EMR	TensorFlow Versão	Componentes instalados com TensorFlow
emr-5.27.1	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.27.0	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.26.0	1.13.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Rótulo de versão do Amazon EMR	TensorFlow Versão	Componentes instalados com TensorFlow
emr-5.25.0	1.13.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.24.1	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.24.0	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Rótulo de versão do Amazon EMR	TensorFlow Versão	Componentes instalados com TensorFlow
emr-5.23.1	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.23.0	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.22.0	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Rótulo de versão do Amazon EMR	TensorFlow Versão	Componentes instalados com TensorFlow
emr-5.21.2	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.21.1	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.21.0	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Rótulo de versão do Amazon EMR	TensorFlow Versão	Componentes instalados com TensorFlow
emr-5.20.1	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.20.0	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.19.1	1.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Rótulo de versão do Amazon EMR	TensorFlow Versão	Componentes instalados com TensorFlow
emr-5.19.0	1.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.18.1	1.9.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.18.0	1.9.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Rótulo de versão do Amazon EMR	TensorFlow Versão	Componentes instalados com TensorFlow
emr-5.17.2	1.9.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.17.1	1.9.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.17.0	1.9.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Apache Tez

O Apache Tez é uma estrutura que cria um grafo acíclico dirigido (DAG) complexo de tarefas para o processamento de dados. Você pode usá-lo como uma alternativa ao Hadoop MapReduce para alguns casos de uso. Por exemplo, você pode executar fluxos de trabalho do Pig e do Hive com o Hadoop MapReduce ou usar o Tez como um mecanismo de execução. Para obter mais informações, consulte https://tez.apache.org/. As versões 4.7.0 e posteriores do Amazon EMR incluem o Tez.

A tabela a seguir lista a versão do Tez incluída na versão mais recente da série 7.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Tez.

Para a versão dos componentes instalados com o Tez nesta versão, consulte Versões de componentes da versão 7.9.0.

Informações sobre a versão Tez para emr-7.9.0

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-7.9.0	Tez 0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

A tabela a seguir lista a versão do Tez incluída na versão mais recente da série 6.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Tez.

Para obter a versão dos componentes instalados com o Tez nessa versão, consulte Release 6.15.0 Component Versions.

Informações de versão do Tez para o emr-6.15.0

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-6.15.0	Tez 0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

A tabela a seguir lista a versão do Tez incluída na versão mais recente da série 5.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Tez.

Para obter a versão dos componentes instalados com o Tez nesse lançamento, consulte Release 5.36.2 Component Versions.

Informações de versão do Tez para o emr-5.36.2

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.36.2	Tez 0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Tópicos

- Criar um cluster com o Tez
- Configurar o Tez
- Interface do usuário da Web do Tez
- Timeline Server
- Histórico de versões do Tez

Criar um cluster com o Tez

Para instalar o Tez, escolha o Apache Tez como uma aplicação quando criar o cluster.

Para criar um cluster com o Tez instalado usando o console

1. Navegue até o novo console do Amazon EMR e selecione Alternar para o console antigo na navegação lateral. Para obter mais informações sobre o que esperar ao alternar para o console antigo, consulte Usar o console antigo.
2. Escolha Create cluster (Criar cluster), Go to advanced options (Ir para opções avançadas).
3. Em Configuração de software, selecione uma Versão do emr-4.7.0 ou versões posteriores.
4. Selecione o Tez com outras aplicações que você deseja que o Amazon EMR instale.
5. Selecione outras opções conforme necessário e escolha Create cluster (Criar cluster).

Para criar um cluster com o Tez usando o AWS CLI

- Use o comando `create-cluster` com a opção `-- applications` para especificar o Tez. O exemplo a seguir cria um cluster com o Tez instalado.

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr create-cluster --name "Cluster with Tez" --release-label emr-7.9.0 \
--applications Name=Tez --ec2-attributes KeyName=myKey \
```

```
--instance-type m5.xlarge --instance-count 3 --use-default-roles
```

Configurar o Tez

Você pode personalizar o Tez definindo valores com a classificação de configuração `tez-site`, que define as configurações no arquivo de configuração `tez-site.xml`. Para obter mais informações, consulte a TezConfigurationdocumentação do Apache Tez. Para alterar o Hive ou o Pig para usar o mecanismo de execução do Tez, use o `hive-site` e classificações de configuração de `pig-properties`, conforme apropriado. Os exemplos são mostrados abaixo.

Exemplo de configuração

Example Exemplo: personalizar o nível de registro raiz do Tez e configurar o Tez como mecanismo de execução para o Hive e o Pig

O comando de exemplo `create-cluster` mostrado a seguir cria um cluster com o Tez, o Hive e o Pig instalados. O comando faz referência a um arquivo armazenado no Amazon S3, `myConfig.json`, que especifica as propriedades da classificação `tez-site` que define `tez.am.log.level` como `DEBUG` e define o mecanismo de execução do Tez para Hive e Pig usando as classificações de configuração `hive-site` e `pig-properties`.

> **Note**
>
> Os caracteres de continuação de linha do Linux (\) são incluídos para facilitar a leitura. Eles podem ser removidos ou usados em comandos do Linux. No Windows, remova-os ou substitua-os por um sinal de interpolação (^).

```
aws emr create-cluster --release-label emr-7.9.0 \
--applications Name=Tez Name=Hive Name=Pig --ec2-attributes KeyName=myKey \
--instance-type m5.xlarge --instance-count 3 \
--configurations https://s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myConfig.json --
use-default-roles
```

Os conteúdos de exemplo de `myConfig.json` são mostrados abaixo.

```
[
    {
```

```
    "Classification": "tez-site",
    "Properties": {
      "tez.am.log.level": "DEBUG"
    }
  },
  {
    "Classification": "hive-site",
    "Properties": {
      "hive.execution.engine": "tez"
    }
  },
  {
    "Classification": "pig-properties",
    "Properties": {
      "exectype": "tez"
    }
  }
]
```

> **Note**
>
> Com as versões 5.21.0 e posteriores do Amazon EMR, você pode substituir as configurações de cluster e especificar classificações de configuração adicionais para cada grupo de instâncias em um cluster em execução. Você faz isso usando o console do Amazon EMR, o AWS Command Line Interface (AWS CLI) ou o AWS SDK. Para obter mais informações, consulte Supplying a Configuration for an Instance Group in a Running Cluster.

Abertura de divisão assíncrona do Tez

Quando há um grande número de arquivos pequenos no caminho da tabela e uma consulta tenta ler todos eles, cada arquivo pequeno que corresponde a cada divisão individual é combinado em uma divisão agrupada do Tez. Então, um único mapper processa a única divisão agrupada do Tez. Como a execução é síncrona, cada divisão individual sob a divisão agrupada é processada uma a uma. Isso exige que os objetos de `RecordReader` processem as divisões de forma síncrona.

O Amazon EMR 6.15.0 introduz configurações que você pode especificar para abrir de forma assíncrona as divisões de entrada em uma divisão agrupada Tez. O recurso foi iniciado pelo TEZ-4397, mas teve regressões no Hive do OSS. O EMR Hive corrigiu as regressões e bugs adicionais na tabela ACID do Hive. Essa melhoria resulta em um desempenho mais rápido das

consultas de leitura quando há um grande número de divisões de entrada em uma única divisão agrupada em Tez.

Name	Classificação	Descrição
`tez.grouping.split.init.threads`	`tez-site`	Especifica o número de threads do daemon que o Tez usa para pré-iniciar o `RecordReaders` e abrir as divisões. Em tabelas ACID, o valor máximo compatível de `tez.grouping.split.init.threads` é 1.
`tez.grouping.split.init.recordreaders`	`tez-site`	Especifica o número de `RecordReaders` a serem mantidos pré-inicializados por threads do daemon. Isso pode ajudar quando a divisão agrupada do Tez contém um grande número de `InputSplits`. A inicialização de `RecordReaders` para processar essas divisões de entrada pode ser feita de forma assíncrona com threads do daemon em vez de processamento sequencial.

Notas de configuração:

Assunto	Detalhes	
Configurações recomendadas	É recomendável definir as configurações acima para os	

Assunto	Detalhes
	valores desejados em ambos `hive-site` `tez-site` e.
Valores correspondentes	Os valores das configurações devem ser os mesmos em ambos `hive-site` `tez-site` e.
Recomendação do LLAP	Não é recomendável usar esse recurso quando o LLAP está habilitado.

Avaliação comparativa de abertura de divisão assíncrona do Tez

Usamos os ambientes e configurações a seguir para comparar a capacidade de abertura de divisão assíncrona do Tez.

- Ambiente de avaliação comparativa: cluster do Amazon EMR com 1 nó primário que usa m5.16xlarge e 16 nós centrais que usam m5.16xlarge.
- Configurações de avaliação comparativa: para simular o cenário de avaliação comparativa em que um grande número de divisões de entrada estão em uma única divisão agrupada do Tez, `tez.grouping.split-count` é definido como `1`.
- Tabela usada para avaliação comparativa: a tabela contém 200 partições, com cada partição contendo um único arquivo. A avaliação comparativa é feita para quando essa tabela contiver arquivos CSV e quando essa tabela contiver arquivos Parquet. Consulta do Hive para avaliação comparativa: `SELECT COUNT(*)` da tabela dez vezes e calcule o runtime médio.
- Configurações para habilitar a abertura de divisão assíncrona do Tez: da seguinte forma:
  - `tez.grouping.split.init.threads` = `4`
  - `tez.grouping.split.init.recordreaders` = `10`

Conjunto de dados	Recurso desabilitado (linha de base)	Recurso habilitado	Melhoria
Conjunto de dados CSV	90.26 segundos	79.20 segundos	12,25%
Conjunto de dados Parquet	54.67 segundos	42.23 segundos	22,75%

Interface do usuário da Web do Tez

O Tez tem sua própria interface de usuário Web. Para visualizar a interface do usuário da Web, consulte o URL a seguir.

```
http://masterDNS:8080/tez-ui
```

Para habilitar a guia Consultas do Hive na interface do usuário da Web do Tez, defina a configuração a seguir.

```
[
  {
    "Classification": "hive-site",
    "Properties": {
      "hive.exec.pre.hooks": "org.apache.hadoop.hive.ql.hooks.ATSHook",
      "hive.exec.post.hooks": "org.apache.hadoop.hive.ql.hooks.ATSHook",
      "hive.exec.failure.hooks": "org.apache.hadoop.hive.ql.hooks.ATSHook"
    }
  }
]
```

Você também pode visualizar detalhes da interface de usuário do Tez, do Spark e da aplicação YARN usando a guia Interfaces de usuário da aplicação da página de detalhes de um cluster no console. As interfaces de usuário (IUs) das aplicações do Amazon EMR são hospedadas fora do cluster e estão disponíveis após o encerramento do cluster. Elas não exigem que você configure uma conexão SSH ou proxy da Web, facilitando a solução de problemas e a análise de trabalhos ativos e do histórico de trabalhos.

Para obter mais informações, consulte Visualizar histórico de aplicações no Guia de gerenciamento do Amazon EMR.

Timeline Server

O servidor de linha do tempo YARN está configurado para execução quando o Tez é instalado. Para visualizar trabalhos enviados por meio do Tez ou de mecanismos de MapReduce execução usando o Timeline Server, visualize a interface do usuário da web usando o URL. `http://master-public-DNS:8188` Para obter mais informações, consulte Visualizar interfaces da Web hospedadas em clusters do Amazon EMR no Guia de gerenciamento do Amazon EMR.

Histórico de versões do Tez

A tabela a seguir informa a versão do Tez incluída em cada versão do Amazon EMR, além dos componentes instalados com a aplicação. Para obter as versões dos componentes em cada versão, consulte a seção Versão de componentes da sua versão em Versões 7.x do Amazon EMR, Versões de lançamento 6.x do Amazon EMR ou Versões de lançamento 5.x do Amazon EMR.

Informações de versão do Tez

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-7.9.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-7.8.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
		namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-7.7.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-7.6.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-7.5.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-7.4.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-7.3.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-7.2.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-5.36.2	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-7.1.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-7.0.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.15.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.14.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-6.13.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.12.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.11.1	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-6.11.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.10.1	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.10.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-6.9.1	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.9.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.8.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-6.8.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.7.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.36.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.36.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.6.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.35.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-6.5.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.4.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.3.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-6.3.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.2.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.2.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-6.1.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.1.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.0.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-6.0.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.34.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.33.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.33.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.32.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.32.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.31.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.31.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.30.2	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.30.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.30.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.29.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.28.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.28.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.27.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.27.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.26.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.25.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.24.1	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.24.0	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.23.1	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.23.0	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.22.0	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.21.2	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.21.1	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.21.0	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.20.1	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.20.0	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.19.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.19.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.18.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.18.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.17.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.17.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.17.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.16.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.16.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.15.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.15.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.14.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.14.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.14.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.13.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.13.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.12.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.12.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.12.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.12.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.11.4	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.11.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.11.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.11.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.11.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.10.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.10.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.9.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.9.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.8.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.8.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.8.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.8.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.7.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.7.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.6.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.6.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.5.4	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.5.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.5.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.5.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.5.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.4.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.4.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.3.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.3.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.3.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.2.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.2.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.2.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.2.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.1.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.1.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.0.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.0.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.0.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-5.0.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.9.6	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.9.5	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-4.9.4	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.9.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.9.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-4.9.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.8.5	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.8.4	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-4.8.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.8.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.8.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-4.8.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.7.4	0.8.3	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.7.3	0.8.3	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rótulo de versão do Amazon EMR	Versão do Tez	Componentes instalados com o Tez
emr-4.7.2	0.8.3	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.7.1	0.8.3	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.7.0	0.8.3	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Notas da versão do Tez por versão

Tópicos

- Amazon EMR 7.8.0 - Notas de lançamento do Tez
- Amazon EMR 7.7.0 - Notas de lançamento do Tez
- Amazon EMR 7.6.0 - Notas de lançamento do Tez
- Amazon EMR 7.5.0 - Notas de lançamento do Tez
- Amazon EMR 7.4.0 - Notas de lançamento do Tez
- Amazon EMR 7.2.0 - Notas de lançamento do Tez
- Amazon EMR 7.1.0: notas da versão do Tez
- Amazon EMR 7.0.0 - Notas de lançamento do Tez
- Amazon EMR 6.15.0: notas de versão do Tez
- Amazon EMR 6.14.0 – Notas da versão do Tez
- Amazon EMR 6.13.0 - Notas da versão do Tez
- Amazon EMR 6.12.0 - Notas da versão do Tez
- Amazon EMR 6.11.0 - Notas da versão do Tez
- Amazon EMR 6.10.0 - Notas da versão do Tez
- Amazon EMR 6.9.0 - Notas da versão do Tez
- Amazon EMR 6.8.0 - Notas da versão do Tez
- Amazon EMR 6.7.0 - Notas da versão do Tez
- Amazon EMR 6.6.0 - Notas da versão do Tez

Amazon EMR 7.8.0 - Notas de lançamento do Tez

Amazon EMR 7.8.0 - Mudanças no chá

Tipo	Descrição
Melhoria	Versão atualizada do bouncycastle para 1.78 para combinar com o hadoop

Amazon EMR 7.7.0 - Notas de lançamento do Tez

Amazon EMR 7.7.0 - Mudanças no Tez

Tipo	Descrição
Melhoria	A `tez.task.relaxed.locality` propriedade no Apache Tez controla se o agendamento de tarefas segue estritamente as restrições de localidade dos dados (localida de do rack e do nó). Quando definido como `true` (padrão no EMR-7.6+), o Tez não impõe a localidade, permitindo que as tarefas sejam atribuídas a qualquer contêiner disponível, o que melhora a utilização dos recursos e reduz os tempos de espera em clusters ocupados.

Amazon EMR 7.6.0 - Notas de lançamento do Tez

Amazon EMR 7.6.0 - Mudanças no Tez

Tipo	Descrição
Melhoria	As restrições de localidade do rack e do nó não serão consideradas ao solicitar o contêiner para uma tarefa, alterando o padrão de para false `tez.task.relaxed.locality`
Melhoria	Ajuste as configurações para desativar o atraso devido à localidade e permitir um retorno não local
Melhoria	TEZ-4547: Adicione o Tez AM JobID ao JobConf

Amazon EMR 7.5.0 - Notas de lançamento do Tez

Amazon EMR 7.5.0 - Mudanças no chá

Tipo	Descrição
Melhoria	Ativou a preempção da tarefa por padrão
Melhoria	Aumento do número de leitores de registros para iniciar de forma assíncrona e proativa até 10
Melhoria	TEZ-4580: Preempção lenta de novos contêineres quando a reutilização está ativada

Amazon EMR 7.4.0 - Notas de lançamento do Tez

Amazon EMR 7.4.0 - Mudanças no chá

Tipo	Descrição
Melhoria	Remova o EventCounter anexador do Tez removido no Hadoop 3.4.0 via HADOOP-17524
Upgrade	Atualize commons-cli para 1.5.0
Upgrade	Atualize a versão Bouncy-castle para 1.70.0

Amazon EMR 7.2.0 - Notas de lançamento do Tez

Amazon EMR 7.2.0 - Mudanças no Tez

Tipo	Descrição
Melhoria	Opção fornecida para desativar as restrições de localidade ao solicitar contêineres para tarefas.

Tipo	Descrição
Correção de bugs	Torne o TaskDependencies# addTaskDependency thread seguro
Upgrade	Atualize a versão TLS para 1.3 para Tez.

Amazon EMR 7.2.0 - Novas configurações

Classificação	Name	Padrão	Descrição
tez-site	tez.task.relaxed.localidade	false	Quando ativadas, as restrições de localidade do rack e do nó não são consideradas ao solicitar um contêiner para uma tarefa.

Amazon EMR 7.1.0: notas da versão do Tez

Amazon EMR 7.1.0: alterações no Tez

Tipo	Descrição
Correção de bugs	TEZ-4394 — Netty4 ShuffleHandler: deve usar 1 thread principal (#256)
Upgrade	Atualiza o Netty para 4.1.100.Final.
Upgrade	Atualiza o Jetty para 9.4.53.v20231009.

Amazon EMR 7.0.0 - Notas de lançamento do Tez

O Amazon EMR 7.0.0 Tez contém todas as alterações e atualizações feitas até e incluindo o Amazon EMR-6.15.0 Tez.

Amazon EMR 7.0.0 - Mudanças no chá

Tipo	Descrição
Upgrade	O Tez Runtime agora usa Java 17 por padrão. Consulte o Guia de versão do EMR 7.0.0 para obter mais detalhes.

Amazon EMR 6.15.0: notas de versão do Tez

Amazon EMR 6.15.0: alterações no Tez

Tipo	Descrição
Recurso	TEZ-4397: abre as divisões de entrada do Tez de forma assíncrona
Upgrade	TEZ-4493: atualiza o Apache Hadoop para 3.3.6

Amazon EMR 6.15.0: recursos do Tez

- Abertura de divisão assíncrona do Tez: o Amazon EMR 6.15.0 introduz configurações que você pode especificar para abrir de forma assíncrona as divisões de entrada em uma divisão agrupada do Tez. O recurso foi iniciado pelo TEZ-4397, mas teve regressões no Hive do OSS. O Hive para Amazon EMR corrigiu as regressões e outros bugs na tabela ACID do Hive. Essa melhoria resulta em uma performance mais rápida das consultas de leitura quando há um grande número de divisões de entrada em uma única divisão agrupada do Tez. Para obter mais informações, consulte Abertura de divisão assíncrona do Tez.

Amazon EMR 6.14.0 – Notas da versão do Tez

Amazon EMR 6.14.0 – Alterações no Tez

Tipo	Descrição
Melhoria	Atualize a versão TLS no Tez para 1.2

Amazon EMR 6.13.0 - Notas da versão do Tez

Amazon EMR 6.13.0: alterações no Tez

Tipo	Descrição
Correção de bugs	REVERT TEZ-4295: não foi possível descompactar os dados. O comprimento do buffer é muito pequeno.
Correção de bugs	REVERT TEZ-4302: `NullPointerException` em `CodecUtils` com `GzipCodec` .
Correção de bugs	REVERT TEZ-4234: o compressor pode causar `IllegalArgumentException` em `Buffer.limit` em que o limite excede a capacidade.
Correção de bugs	REVERT TEZ-4135: Melhore a alocação de memória ao executar leituras na memória.

Amazon EMR 6.12.0 - Notas da versão do Tez

Amazon EMR 6.12.0 - Alterações no Tez

Tipo	Descrição
Melhoria	Adicionado suporte para runtime do JDK 11 e do JDK 17

Tipo	Descrição
Correção de bugs	TEZ-4492: atualize o Bowerrc para usar o espelho bower.herokuapp para evitar o problema CERT_EXPIRE do registro do Bower (BOWER-2608)
Upgrade	Atualizado o Surefire para 3.0.0-M7

Amazon EMR 6.11.0 - Notas da versão do Tez

Amazon EMR 6.11.0 - Alterações no Tez

Tipo	Descrição
Bug	Corrigida a transição de estado de vértices inválida durante a limpeza em nível de vértice de dados embaralhados
Bug	Corrigida a limpeza em nível de DAG ou vértice de dados embaralhados que não funcionam
Melhoria	Ative tez.am.dag.cleanup.on.completion por padrão para limpar os dados aleatórios de concluídos DAGs

Amazon EMR 6.10.0 - Notas da versão do Tez

Amazon EMR 6.10.0 - Alterações no Tez

Tipo	Descrição
Recurso	Habilitar `tez.runtime.transfer.data-via-events.enabled` por padrão
Backport	TEZ-4450: corrige a falha na busca de dados embaralhados quando os dados embaralha

Tipo	Descrição
	dos são transferidos por meio de eventos de movimentação de dados
Backport	TEZ-4460: corrige o erro de tempo limite de leitura na busca de dados embaralhados do Tez Shuffle Handler
Backport	TEZ-4455: Adicione um ShuffleHandler pipeline para LoggingHandler melhor depuração
Bug	Corrija a tarefa do Tez que fica presa de forma intermitente quando a preempção da tarefa está habilitada

Amazon EMR 6.9.0 - Notas da versão do Tez

Amazon EMR 6.9.0 - Alterações no Tez

Tipo	Descrição
Upgrade	O Tez é atualizado para 0.10.2. Para obter mais informações, consulte o log de alterações do Apache Tez 0.10.2.
Upgrade	Atualizar o Hadoop para 3.3.3.
Bug	Desabilite `tez.runtime.transfer.data-via-events.enabled` por padrão devido ao TEZ-4450.

Amazon EMR 6.8.0 - Notas da versão do Tez

Amazon EMR 6.8.0 - Alterações no Tez

Tipo	Descrição
Backport	TEZ-3363: excluir dados intermediários em nível de vértice para o Shuffle Handler
Backport	TEZ-4129: excluir dados de tentativas intermediárias para tentativas malsucedidas do Shuffle Handler
Backport	TEZ-4430: corrigida a propriedade tez.task.launch.cmd-opts corrigida que não funciona

Amazon EMR 6.7.0 - Notas da versão do Tez

Amazon EMR 6.7.0 - Alterações no Tez

Tipo	Descrição
Backport	TEZ-4403: Atualize SLF4 a versão J para 1.7.36
Backport	TEZ-4405: substituir log4j 1.x por reload4j
Backport	TEZ-4411: Falha na compilação do Tez: .js não encontrado FileSaver

Amazon EMR 6.6.0 - Notas da versão do Tez

Amazon EMR 6.6.0 - Alterações no Tez

Tipo	Descrição
Backport	TEZ-3918: corrigida a propriedade tez.task.log.level que não funciona.
Backport	TEZ-4353: atualize commons-io para 2.8.0.
Backport	TEZ-4114: remova do tez a dependência direta do jetty.
Backport	TEZ-4323: os arquivos jar do Jetty foram removidos do pacote do distribuidor com o TEZ-4114.

Trino

O Trino é um mecanismo de consulta de código aberto projetado para consultas interativas em uma ampla variedade de fontes de dados. Isso pode incluir bancos de dados relacionais, dados baseados em arquivos, dados HDFS e outros. O objetivo mais comum do Trino com o Amazon EMR é executar consultas SQL complexas em grandes conjuntos de dados armazenados no Amazon S3. Também é compatível com o ANSI SQL, o que o torna familiar para engenheiros de banco de dados, analistas de dados e cientistas de dados que estão familiarizados com o SQL.

> **Note**
>
> O PrestoSQL foi renomeado para Trino em dezembro de 2020. As versões 6.4.0 e posteriores do Amazon EMR geralmente se referem ao Trino, enquanto as versões anteriores se referem ao PrestoSQL.

> **Important**
>
> O PrestoSQL, a versão anterior do Trino, ainda está disponível para uso com o Amazon EMR. No entanto, é altamente recomendável usar o Trino no futuro com o Amazon EMR. Observe também que o Trino e o PrestoSQL não podem ser executados simultaneamente no mesmo cluster.

A tabela a seguir lista a versão do Trino incluída na versão mais recente do Amazon EMR, além dos componentes que o Amazon EMR instala com o Trino. Para obter a versão dos componentes instalados com o Trino nessa versão, consulte Release 7.9.0 Component Versions.

Informações de versão do Trino para o emr-7.9.0

Rótulo de versão do Amazon EMR	Trino	Componentes instalados com o Trino
emr-7.9.0	Trino 467	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-

Rótulo de versão do Amazon EMR	Trino	Componentes instalados com o Trino
		server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-trino, hcatalog-server, mariadb-server, trino-coordinator, trino-worker

Tópicos

- História e design de Trino
- Conceitos de Trino
- Configurar o Trino no Amazon EMR
- Melhores práticas para Trino no Amazon EMR
- Trino
- Histórico de versões do Trino

História e design de Trino

A Trino é especializada em consultar grandes conjuntos de dados de várias fontes diferentes. O Trino pode acessar e consultar o HDFS em um caso de uso tradicional de big data, mas também pode consultar fontes adicionais, como bancos de dados relacionais e bancos de dados NoSQL. O Trino começou originalmente como um fork do mecanismo de consulta Presto, em 2019. Desde então, ele foi desenvolvido independentemente da base de código do Presto.

Para obter mais informações sobre o mecanismo de consulta Trino e como ele é usado, consulte o site do Trino. Para ler a documentação do código-fonte do Trino, consulte Visão geral do Trino.

Conceitos de arquitetura

O Trino pode executar consultas rápidas e eficientes porque processa dados paralelamente em um cluster. Ele foi projetado pensando na consulta de um data lake, pois é especializado em consultas em grandes volumes de dados, normalmente em casos de uso envolvendo Hadoop e HDFS. Mas

também pode consultar bancos de dados relacionais tradicionais. Para obter mais informações, consulte Arquitetura na documentação do Trino.

Componentes do Trino

O Trino tem alguns componentes principais de arquitetura que funcionam juntos para agilizar a execução das consultas. É útil ter um conhecimento prático deles quando você ajusta seu cluster para obter um melhor desempenho:

- O coordenador é responsável pela orquestração de consultas. Ele analisa e otimiza as consultas SQL recebidas, gera planos de execução, atribui tarefas aos nós de trabalho e coleta e monta os resultados da consulta. Além disso, ele monitora o uso de recursos e rastreia o status dos nós de trabalho. Para obter mais informações, consulte Coordenador na documentação do Trino.
- Os nós de trabalho lidam com o processamento de dados para consultas. Depois que o coordenador atribui tarefas, os trabalhadores recuperam dados, realizam as operações necessárias, como uniões e agregações, e trocam dados intermediários com outros trabalhadores. Para obter mais informações, consulte Worker na documentação do Trino.
- Os conectores são plug-ins que permitem que o Trino se conecte e consulte várias fontes de dados. Cada conector sabe como acessar e recuperar dados de sua fonte, como Amazon S3, Apache Hive ou bancos de dados relacionais. Esses conectores mapeiam os dados de origem para a estrutura do esquema de Trino.
- Um catálogo é uma coleção lógica de esquemas e tabelas associados a um conector específico. Definidos no coordenador, os catálogos permitem que o Trino trate diferentes fontes de dados como um único namespace. Isso faz com que os usuários possam consultar várias fontes juntas, como Hive e MySQL, de forma unificada na mesma consulta.
- Clientes como o Trino CLI se conectam por meio de drivers JDBC e ODBC ao coordenador do Trino para enviar consultas SQL. O coordenador gerencia o ciclo de vida da consulta, fornecendo resultados ao cliente para análises ou relatórios adicionais.

Executar consultas

Para entender como o Trino pega instruções SQL e as executa como consultas, consulte Conceitos do Trino na documentação do Trino.

Conceitos de Trino

Os procedimentos nesta seção mostram como configurar um cluster do Amazon EMR para consultar as fontes de dados do metastore usando o Trino. Esses metastores, que incluem o AWS Glue Data Catalog, armazenam metadados e objetos de banco de dados e gerenciam permissões de acesso. Os procedimentos abrangem pré-requisitos, configurações recomendadas, criação de conectores e execução de consultas em tabelas do metastore.

Tópicos

- Conclua as etapas de pré-requisito para usar o Amazon EMR com Trino
- Inicialização de um cluster do Amazon EMR com o Trino
- Conectar-se ao nó primário do cluster do Amazon EMR

Conclua as etapas de pré-requisito para usar o Amazon EMR com Trino

Se você não usou ou não criou um cluster do Amazon EMR AWS, conclua essas etapas de pré-requisito antes de criar um cluster do Amazon EMR com o Trino.

AWS configuração do ambiente

Conclua estas etapas para configurar sua AWS conta, caso ainda não tenha feito isso:

1. Inscreva-se em uma AWS conta, se ainda não tiver uma. Para obter mais informações, consulte Criar uma AWSAWS conta no Guia de referência de gerenciamento de contas.
2. Faça login em sua conta da como usuário administrador.
3. Crie um grupo e atribua usuários a ele.
4. Crie um par de EC2 chaves da Amazon, que você pode usar posteriormente para proteger a comunicação entre recursos com SSH. Essa etapa é necessária se você planeja se conectar ao nó primário para realizar tarefas. Para obter mais informações, consulte Connect to the the primary node to the primary node using SSH do cluster do Amazon EMR usando SSH.

Inicialização de um cluster do Amazon EMR com o Trino

A seguir, descrevemos as opções de configuração corretas ao criar um cluster usando o Trino.

Usando um conector Hive para disponibilizar dados para consulta

Você pode configurar um conector Trino para um metastore do Hive com o objetivo de consultar dados do metastore do seu cluster. Um metastore é uma camada de abstração que disponibiliza conteúdo ou dados baseados em arquivos como tabelas, facilitando a consulta. Você precisa configurar um conector no Amazon EMR para disponibilizar as tabelas de metastore do Hive para o cluster. O procedimento a seguir mostra como fazer isso:

1. Escolha AWS Glue no console e crie uma tabela com base em seus dados de origem no Amazon S3. Uma tabela no AWS Glue Data Nesse contexto, faz sentido criar a tabela manualmente, criando colunas conforme desejar, a partir dos dados de origem. Para obter mais informações sobre a criação de tabelas no AWS Glue a partir de dados semiestruturados no Amazon S3, consulte Criação de tabelas usando o console no Guia do usuário do AWS Glue.
2. Defina sua configuração como parte da criação do cluster. Selecione a guia Configuração. As configurações são especificações opcionais para o cluster. Ao inserir uma configuração, adicione JSON como no exemplo a seguir, que instrui Trino a usar o AWS Glue Data Catalog como seu metastore externo do Hive para metadados de tabelas:

   ```
   {
       "classification": "trino-connector-hive",
       "properties": {
           "hive.metastore": "glue"
       }
   }
   ```

   Como alternativa, você pode aplicar configurações na seção Configurações de software ao criar um cluster.

   Além disso, você pode configurar outros tipos de conectores, como para conexão com o Apache Iceberg. Para obter mais informações, consulte Usar um cluster Iceberg com Trino no Guia de lançamento do Amazon EMR. Definir configurações adicionais é opcional.

Para continuar com as etapas iniciais, consulte. Conectar-se ao nó primário do cluster do Amazon EMR

Crie um cluster com o Trino

A seguir, descrevemos as opções de configuração corretas ao criar um cluster que você deseja usar com o Trino.

> **Important**
>
> Antes de criar seu cluster, conclua a configuração do AWS Glue Data Catalog como seu metastore do Hive, o que recomendamos para começar. Para obter mais informações, consulte Usando um conector Hive para disponibilizar dados para consulta.

1. No AWS console, selecione Amazon EMR nos serviços. Quando você escolhe o Amazon EMR, se você tiver clusters existentes, seu EMR em EC2 clusters é listado.
2. Selecione Criar cluster. A partir daqui, você inicia o processo de criação de um cluster.
3. Dê um nome ao seu cluster e escolha uma versão do Amazon EMR. Você pode escolher a versão mais atual para o tutorial.
4. Escolha o pacote Trino, que tem o aplicativo Trino pré-selecionado. Os pacotes são configurados para facilitar quando você sabe a finalidade do cluster com antecedência. Caso contrário, você pode simplesmente marcar a caixa de seleção para Trino.
5. Para configuração de cluster, escolha Uniform instance groups. Vá em frente e remova outros grupos de instâncias.
6. Escolha um tipo de instância. Geralmente, recomendamos que você escolha um tipo de instância com pelo menos 16 GiB de memória. Além disso, para escalabilidade e provisionamento de clusters, escolha Definir tamanho do cluster manualmente.
7. Neste ponto, defina a configuração da metastore do Hive para apontar para Glue. AWS Isso está detalhado na seçãoUsando um conector Hive para disponibilizar dados para consulta. Conclua isso antes de criar o cluster.
8. Selecione Criar cluster. Pode demorar alguns minutos para terminar.

   As etapas aqui não abrangem todas as etapas de configuração em detalhes. Mais informações sobre a configuração de um cluster estão disponíveis em Planejar, configurar e iniciar clusters do Amazon EMR.

> **Note**
>
> Não selecione Presto e Trino para uso no mesmo cluster. Não há suporte para executá-los juntos. Também é recomendável que, se você executar o Trino, não execute nenhum outro aplicativo no cluster, como o Spark.

Conectar-se ao nó primário do cluster do Amazon EMR

Provisione dados de teste e configure permissões

Você pode testar o Amazon EMR com o Trino usando o AWS Glue Data Catalog e seu metastore Hive. Essas etapas de pré-requisito descrevem como configurar os dados de teste, caso ainda não tenha feito isso:

1. Crie uma chave SSH para usá-la para criptografia de comunicação, caso ainda não tenha criado.
2. Você pode escolher entre vários sistemas de arquivos para armazenar dados e arquivos de log. Para começar, crie um bucket do Amazon S3. Dê ao bucket um nome exclusivo. Ao criá-lo, especifique a chave de criptografia que você criou.

   > **Note**
   >
   > Escolha a mesma região para criar seu bucket de armazenamento e o cluster do Amazon EMR.

3. Escolha o bucket que você criou. Escolha Criar pasta e dê à pasta um nome memorável. Ao criar a pasta, escolha uma configuração de segurança. Você pode escolher as configurações de segurança para os pais ou torná-las mais especializadas.
4. Adicione dados de teste à sua pasta. Para os fins deste tutorial, usar um.csv de registros separados por vírgula funciona bem para concluir esse caso de uso.
5. Depois de adicionar dados a um bucket do Amazon S3, configure uma tabela no AWS Glue para fornecer uma camada de abstração para consultar os dados.

Conecte-se e execute consultas

A seguir, descrevemos como você se conecta e executa consultas em um cluster executando o Trino. Antes de fazer isso, certifique-se de configurar o conector do Hive metastore, descrito no procedimento anterior, para que as tabelas do metastore fiquem visíveis.

1. Recomendamos usar o EC2 Instance Connect para se conectar ao seu cluster, pois ele fornece uma conexão segura. Escolha Connect to the Primary node using SSH no resumo do cluster. A conexão exige que o grupo de segurança tenha uma regra de entrada para permitir conexões

pela porta 22 com clientes na sub-rede. Você também deve usar o usuário hadoop ao se conectar.

2. Inicie a CLI do Trino executando. `trino-cli` Isso permite que você execute comandos e consulte dados com o Trino.
3. Executar `show catalogs;`. Verifique se o catálogo da colmeia está listado. Isso fornece uma lista de catálogos disponíveis, que contêm armazenamentos de dados ou configurações do sistema.
4. Para ver os esquemas disponíveis, execute`show schemas in hive;`. A partir daqui, você pode executar `use` *`schema-name`*`;` e incluir o nome do esquema. Em seguida, você pode executar `show tables;` para listar tabelas.
5. Consulte uma tabela executando um comando como`SELECT * FROM` *`table-name`*, usando o nome de uma tabela em seu esquema. Se você já executou a `USE` instrução para se conectar a um esquema específico, não precisa usar a notação de duas partes, como. *`schema table`*.

Configurar o Trino no Amazon EMR

Tópicos

- Configurando conectores para Trino
- Monitoramento

Configurando conectores para Trino

Conectando-se ao AWS Glue como sua metastore do Hive

É importante e útil entender que você pode configurar o AWS Glue Data Catalog como seu metastore do Hive ao executar consultas com o Trino. Para obter informações adicionais, incluindo etapas para configurar um cluster com um metastore do Hive, consulte Usando o AWS Glue Data Catalog como metastore do Hive.

Para obter informações sobre a integração do EMR no EKS com o AWS Glue, consulte as seguintes melhores práticas: integração de contêineres do EMR com o Glue. AWS

Conectando-se às tabelas do Iceberg ao usar o Trino com o Amazon EMR

O Iceberg é um formato de tabela aberto para tabelas analíticas. Ele foi criado para mecanismos como Spark e Trino consultarem big data das mesmas tabelas, usando consultas SQL. Ele inclui

recursos como isolar leituras e gravações de dados, para que um leitor possa evitar consultar dados parcialmente atualizados, por exemplo. Ele também oferece suporte a recursos de estado, como instantâneos. Ele fornece uma camada de abstração por meio do uso de metadados e arquivos de manifesto. Eles descrevem o esquema da tabela e facilitam a consulta de dados sem precisar saber muitos detalhes sobre como eles são formatados ou organizados. Quando você está conectado, você pode ler dados das tabelas, atualizar dados ou gravar novos dados nos arquivos subjacentes.

Há um workshop disponível que mostra como configurar tabelas Iceberg com o Amazon EMR e o Glue AWS . Para obter mais informações, consulte Workshop de análise - Configurar e usar tabelas Apache Iceberg em seu data lake.

Conectar-se com os clientes

Você pode se conectar ao Trino usando um driver JDBC disponível. Para obter mais informações, consulte o driver JDBC na documentação do Trino.

Monitoramento

Você pode monitorar clusters do Amazon EMR por meio do. AWS Management Console Para obter mais informações, consulte Exibição e monitoramento de um cluster do Amazon EMR enquanto executa trabalhos. O Amazon EMR também envia suas métricas de monitoramento para. Amazon CloudWatch Para obter mais informações sobre como monitorar um cluster do Amazon EMR, consulte Amazon CloudWatch os eventos e métricas do Amazon EMR.

Melhores práticas para Trino no Amazon EMR

A arquitetura do Trino foi projetada para consultas SQL rápidas e distribuídas em grandes conjuntos de dados em várias fontes de dados, seguindo um modelo coordenador-trabalhador, em que cada componente tem uma função especializada na execução de consultas. Há algumas áreas ou categorias nas quais você pode se concentrar para configurar seu cluster Amazon EMR executando o Trino para obter o melhor desempenho. Incluindo o seguinte:

- Ajustando as configurações do cluster para otimização da memória.
- Otimizando as configurações para particionamento e distribuição de dados.
- Usando a filtragem dinâmica para reduzir a contagem de resultados da consulta.

Algumas dessas configurações são ajustadas automaticamente quando você usa o Trino com o Amazon EMR. Outros podem ser configurados manualmente por meio do console ou por meio de

comandos CLI. Os tópicos desta seção ajudam você a configurar os dados e o cluster de forma otimizada.

Tópicos

- Principais áreas de foco para melhoria de desempenho
- Colete e utilize estatísticas de tabelas
- Desafios comuns ao escalar as cargas de trabalho do Trino

Principais áreas de foco para melhoria de desempenho

O Trino maximiza o paralelismo de consultas e a otimização da memória. Essa arquitetura oferece flexibilidade ao permitir que ela consulte várias fontes de dados variadas e, ao mesmo tempo, escale com eficiência. As principais áreas de melhoria de desempenho no Trino incluem as listadas abaixo.

Otimização de memória

O gerenciamento de memória no Trino é fundamental para alcançar alto desempenho e estabilidade, especialmente quando você executa consultas grandes e complexas. O Trino usa um modelo de memória distribuída. Nesse modelo, a memória é alocada entre os nós de trabalho para processar tarefas, agregações, uniões e outras operações. A lista a seguir apresenta uma coleção dessas configurações:

- query.max-memory — Define a memória máxima disponível para uma única consulta em todo o cluster. Esse é um limite rígido; se uma consulta exceder essa memória, ela falhará.
- consultar. max-memory-per-node — Define a memória máxima que uma consulta pode consumir em cada nó de trabalho. Essa configuração garante que nenhuma consulta monopolize os recursos de nenhum trabalhador.
- Tamanho do heap da JVM — Configurado no nível da JVM, ele define o tamanho máximo do heap para o processo do servidor Trino em cada nó. Esse valor geralmente deve ser maior do que as configurações relacionadas à memória (essa é a soma da consulta). max-memory-per-nodee memória. heap-headroom-per-node) no Trino para evitar que o sistema fique sem memória no nível da JVM.
- memória. heap-headroom-per-node — Especifica a quantidade de memória do buffer a ser omitida do tamanho do heap da JVM para operações que não sejam de consulta. Isso é crucial para garantir despesas gerais suficientes para operações internas e coleta de lixo.

Filtragem dinâmica

A filtragem dinâmica no Trino é uma técnica de otimização que melhora o desempenho da consulta ao reduzir a quantidade de dados processados, principalmente durante as junções. Ele aplica dinamicamente condições de filtro para limitar os dados digitalizados por um lado de uma junção, com base nos dados vistos no outro lado, o que é especialmente útil em consultas em que um lado da junção é altamente seletivo (o que significa que contém um pequeno subconjunto de dados). Ele é ativado por padrão no Amazon EMR. Veja a seguir um exemplo de consulta:

```
SELECT orders.order_id, orders.total_amount
FROM orders
JOIN customers ON orders.customer_id = customers.customer_id
WHERE customers.country = 'France';
```

Sem a filtragem dinâmica, o Trino verifica toda a tabela de pedidos em uma junção, embora apenas um pequeno subconjunto de clientes (os da França) seja relevante. Essa abordagem lê todas as linhas na tabela de pedidos, resultando em altos custos de E/S e processamento. Com a filtragem dinâmica, o Trino inicialmente verifica a tabela de clientes menores, recupera valores de customer_id somente para clientes da França e, em seguida, aplica esse subconjunto como um filtro nos pedidos. Isso significa que somente as linhas relevantes dos pedidos — aquelas com um customer_id correspondente ao subconjunto filtrado — são digitalizadas, reduzindo significativamente os registros processados.

Derramar

No Trino, o vazamento de disco permite que os resultados da consulta intermediária sejam transferidos para o disco, permitindo que consultas com uso intensivo de memória sejam concluídas, mesmo que excedam os limites de memória definidos por ou. `query_max_memory` `query_max_memory_per_node` Por padrão, o Trino impõe esses limites para garantir uma alocação justa de memória e evitar o impasse do cluster. No entanto, quando uma consulta grande ultrapassa esses limites, ela corre o risco de ser encerrada. O vazamento de disco resolve isso usando`revocable memory`, permitindo que uma consulta empreste memória adicional que pode ser revogada se forem necessários recursos em outro lugar. Quando a memória é revogada, os dados intermediários são transferidos para o disco, permitindo que as consultas continuem sendo processadas sem exceder os limites de memória. Observe que uma consulta que é forçada a ser transferida para o disco pode ter um tempo de execução mais longo e, portanto, está desativada por padrão. Para habilitar o vazamento no Amazon EMR, use a seguinte configuração:

- `spill-enabled=true`— Permite o vazamento de disco quando o uso da memória excede os limites disponíveis.
- `spill-paths`— Define os diretórios onde os dados derramados são armazenados,. `spill-paths=/mnt/spill`

Colete e utilize estatísticas de tabelas

A coleta de estatísticas de tabelas permite que o otimizador baseado em custos da Trino tome decisões informadas sobre pedidos de junção, remoção de filtros e remoção de partições, resultando em melhor desempenho.

Você pode usar o `ANALYZE` comando para coletar estatísticas das tabelas Hive ou Iceberg:

```
ANALYZE sales;
```

Coletar estatísticas em tabelas amplas pode sobrecarregar os recursos. Recomendamos especificar um subconjunto de colunas que são usadas em junções, filtros ou operações de agrupamento.

Esse é outro comando útil. Ele exibe as estatísticas atuais de uma tabela para verificar se as estatísticas estão atualizadas.

```
show stats for table_name;
```

Desafios comuns ao escalar as cargas de trabalho do Trino

Os principais benefícios de usar o Amazon S3 com o Trino são a capacidade do S3 de escalar para grandes volumes de dados e a relação custo-benefício do S3. Mas quando você consulta grandes volumes de dados, uma coleção de problemas de desempenho relacionados pode acontecer de vez em quando. Isso pode resultar da forma como os dados são armazenados, de configurações que restringem o bom desempenho ou de outros motivos. Quando esses problemas ocorrem, há medidas eficazes que você pode tomar para evitá-los ou mitigá-los.

Esta seção começa com uma lista de otimizações gerais que você pode implementar para aumentar o desempenho da consulta em grandes volumes de dados. Depois disso, os problemas comuns são detalhados e as mitigações são fornecidas para cada um.

Este tópico foi extraído da seguinte apresentação da conferência: Acelere o desempenho em grande escala: melhores práticas para Trino com o Amazon S3.

Otimizar o layout de dados para conjuntos de dados grandes

Os gargalos de desempenho não são raros quando você está consultando grandes conjuntos de dados. Mas existem melhores práticas que você pode implementar para ter uma vantagem inicial ao usar o Trino para consultar dados no Amazon S3. Incluindo o seguinte:

- Particionamento — Particionamento significa organizar dados em uma hierarquia e armazenar dados relacionados juntos, com base em atributos relacionados. O particionamento faz com que as consultas não precisem verificar tantos dados irrelevantes e isso resulta em um melhor desempenho da consulta. Você pode usar várias estratégias de particionamento, como organizar os dados de origem com prefixos, especificamente por intervalos de datas, regiões ou outros atributos. Para obter informações mais detalhadas sobre dados de particionamento no Amazon S3 para aumentar o desempenho, consulte a postagem do blog Comece a gerenciar partições para tabelas do Amazon S3 apoiadas AWS pelo Glue Data Catalog ou a postagem Top 10 Performance Tuning Tips for. Amazon Athena
- Armazenamento em compartimentos — O agrupamento de dados relacionados em arquivos comuns. Por exemplo, se você consultar dados de acordo com uma região geográfica, como um estado, poderá aumentar o desempenho da consulta agrupando todos os dados de um determinado estado no mesmo arquivo ou grupo de arquivos. Para que isso funcione melhor, baseie seu agrupamento em um atributo de dados com alta cardinalidade, como um estado ou província, por exemplo. Além disso, você pode levar em consideração seus padrões de consulta. Um exemplo disso pode significar agrupar dados da Califórnia e do Oregon, caso suas consultas normalmente leiam dados desses estados juntos.
- Gerenciamento de prefixos do S3 — Você pode usar os prefixos do Amazon S3 para implementar uma estratégia de particionamento. Se você usar apenas um único prefixo para um bucket do Amazon S3, como uma data específica, por exemplo, isso pode levar a um grande número de solicitações e resultar em um erro HTTP 503. Recomendamos o uso de prefixos para adicionar condições adicionais e organizar seus dados de origem com mais eficiência. Para obter mais informações, consulte Organizar objetos usando prefixos na documentação do Amazon S3. O breve exemplo a seguir mostra um prefixo que resulta em uma melhor taxa de transferência de solicitações:. `s3://bucket/country=US/dt=2024-06-13` Neste exemplo, o país e a data são incluídos no prefixo, o que resulta em menos leituras do que um caso em que o prefixo inclui somente a data.

  A mitigação dos erros HTTP 503 é discutida com mais detalhes na seção de desaceleração do HTTP que segue neste tópico.

- Otimizando o tamanho dos dados — você pode executar o comando OPTIMIZE para definir uma configuração que conduza às consultas com melhor desempenho. Para executá-lo em tabelas externas do Hive, siga estas etapas:
  - Use `OPTIMIZE` com o seguinte parâmetro:`hive.non-managed-table-writes-enabled=true`. Para obter mais informações sobre essa propriedade, consulte Propriedades gerais de configuração do Hive.
  - Defina o seguinte parâmetro de sessão: `SET SESSION` *`catalog`*`.non_transactional_optimize_enabled=true`
  - Execute o `OPTIMIZE` comando:`ALTER TABLE` *`catalog`*`.`*`schema`*`.`*`table`* `EXECUTE optimize(file_size_threshold => '128MB')`. Nesse caso, `file_size_threshold` é de 100 MB por padrão. Aumentar esse limite, conforme mostrado na amostra, fará com que arquivos abaixo de 128 MB sejam mesclados.
- Configurar novas tentativas — Você pode aumentar o limite de novas tentativas, o que pode reduzir a chance de erros HTTP 503, definindo o seguinte:. `s3.max-error-retries` Isso se aplica quando você usa a TrinoFileSystem API e a versão 449 do Trino ou posterior. Por outro lado, no caso em que você está usando o Amazon EMR com Trino, você usa o EMRFS para acessar o Amazon S3. Com o EMRFS, você pode aumentar o número de retiradas alterando o parâmetro. `fs.s3.maxRetries`
- Escolha uma classe de armazenamento do Amazon S3 — Escolher a classe de armazenamento apropriada para dados em diferentes pontos do ciclo de vida pode ajudar tanto no desempenho quanto no custo, com base em seus requisitos para coleções de dados específicas. Para obter mais informações, consulte Entender e gerenciar as classes de armazenamento do Amazon S3 na documentação do Amazon S3.
- Migrar para o Iceberg — Outra solução para mitigar problemas de desempenho, especificamente em relação à execução de consultas em arquivos pequenos, é migrar para tabelas do Iceberg. O Iceberg tem recursos que lidam bem com arquivos pequenos.
- Use compactação automática de dados — Se você utiliza tabelas Iceberg, a compactação automática de dados com o AWS Glue Data Catalog pode otimizar o tamanho dos dados e resultar em melhor desempenho de consultas.

Desafios comuns ao consultar grandes conjuntos de dados

Esta seção lista uma coleção de problemas comuns que podem ocorrer quando você acumula um grande conjunto de dados no Amazon S3 e o consulta com o Trino. Cada seção mostra maneiras de

resolver o problema ou reduzir seu impacto nas consultas. Cada um dos problemas descritos nas seções a seguir foi reproduzido e testado usando um conector Hive.

Grandes varreduras de dados

Quando sua consulta precisa escanear grandes conjuntos de dados, isso pode levar a problemas como baixo desempenho da consulta e maior custo de armazenamento. Grandes volumes de dados podem resultar de um rápido crescimento de dados ou de um planejamento que não resulte na movimentação de dados legados dentro de um prazo adequado. Isso pode gerar consultas mais lentas.

Para reduzir os impactos no desempenho causados pela varredura de grandes conjuntos de dados, recomendamos que você utilize particionamento e armazenamento em intervalos:

- O particionamento agrupa dados relacionados, com base em seus atributos. Usar o particionamento de forma eficaz pode melhorar muito o desempenho das consultas.
- O agrupamento por compartimentos se refere ao agrupamento de dados em arquivos ou compartimentos de acordo com colunas de dados específicas e relacionadas. O armazenamento em compartimentos normalmente significa manter fisicamente os arquivos de dados de origem relacionados juntos.

Para ilustrar como a mitigação pode funcionar para grandes varreduras de dados, suponha que você armazene e consulte dados que tenham registros com um atributo de estado, que pode ser atribuído à Califórnia ou ao Alasca, e esse atributo de estado seja uma de suas condições de consulta. Você pode melhorar o desempenho da consulta armazenando dados de cada estado em um bucket S3 separado ou particionando seus dados com base no estado, usando um prefixo S3. Esse particionamento e compartimentação também podem melhorar o desempenho se você os basear em uma coluna adicional, como um atributo de data, por exemplo.

> **Note**
>
> Se uma coluna tiver alta cardinalidade e você quiser usá-la para agrupar dados, recomendamos o uso de compartimentos nesse caso. Por outro lado, geralmente, as chaves de partição devem ter menor cardinalidade.

Usando vários tipos de armazenamento S3

Geralmente, você escolhe os tipos de armazenamento com base nos requisitos de desempenho, acesso a dados, resiliência e custo para suas workloads. Pode haver compensações entre custo e desempenho. É importante escolher a classe de armazenamento apropriada do Amazon S3 que corresponda aos seus padrões de acesso aos dados. Há dois padrões principais de acesso:

- Dados que são acessados de forma conhecida ou previsível. Geralmente, se você tem dados que são acessados com pouca frequência, o S3 Standard IA pode ser uma boa escolha, pois ajuda a reduzir custos. Se você acessa dados com frequência, o S3 Standard é melhor para acessar com o Amazon EMR e o Trino.
- Dados que são acessados de forma desconhecida ou imprevisível. Isso pode exigir o uso de outras classes de armazenamento do Amazon S3. Há vantagens e desvantagens entre as classes de armazenamento do Amazon S3. Isso inclui latência, custo e disponibilidade de armazenamento. Você pode escolher um tipo de armazenamento S3 apropriado, com base em suas cargas de trabalho e padrões de acesso. Para obter descrições dos benefícios de cada classe, consulte Classes de armazenamento do Amazon S3.

Usando compactação

Você também pode usar a compactação automática do Iceberg, se usar tabelas do Iceberg, o que resulta em tamanhos de arquivo mais ideais, para aumentar a eficiência da consulta. Para obter mais informações, consulte O AWS Glue Data Catalog agora oferece suporte à compactação automática de tabelas Apache Iceberg.

Erros de lentidão de HTTP

Isso ocorre quando a taxa de solicitação excede um limite pré-configurado em um prefixo do Amazon S3. O erro HTTP que ocorre com mais frequência quando esse estado é atingido é o seguinte: Erro 503: reduza sua taxa de solicitações. A origem desse problema pode estar enraizada na presença de um grande número de arquivos pequenos, devido ao número de divisões que devem ser criadas para ler os dados. Há algumas maneiras de mitigar o problema:

- Aumente o limite de novas tentativas para solicitações do Amazon S3 no Trino. Isso é definido para uso do EMRFS `fs.s3.maxretries` no Trino 449.
- Otimize os tamanhos dos arquivos, o que também pode resultar em uma menor taxa de solicitações.

Para obter mais informações sobre como o Trino determina o número de divisões em um conjunto de dados a serem consultados, consulte Propriedades de configuração de ajuste de desempenho na documentação do conector Hive.

Dificuldade em consultar arquivos pequenos

Consultar muitos arquivos pequenos pode resultar em sobrecarga de E/S, devido ao alto número de solicitações GET e LIST, e, consequentemente, afetar negativamente o desempenho da consulta. A otimização do tamanho do arquivo pode melhorar o desempenho da consulta. Há várias maneiras de fazer isso:

- Consolide os dados em menos arquivos maiores. (Geralmente, recomendamos manter os tamanhos dos arquivos em torno de 128 MB.) Você pode fazer isso com ferramentas ao ingerir dados, como em um pipeline de ETL, ou consolidar dados manualmente. Se essas soluções não estiverem disponíveis para você, as opções restantes podem ser mais adequadas para você.
- Execute o comando `OPTIMIZE`.
- Defina o parâmetro `SESSION`.

Observe que o Iceberg tem um recurso disponível para mesclar arquivos pequenos em arquivos maiores, que é a compactação automática. Ele funciona com arquivos gerenciados com o AWS Glue Data Catalog. Para obter mais informações, consulte O AWS Glue Data Catalog agora oferece suporte à compactação automática de tabelas Apache Iceberg.

Consultas que incluem dados que não são necessários

É comum que os dados cresçam, o que torna imperativo rastrear seus padrões de acesso aos dados e movê-los adequadamente à medida que envelhecem ou se tornam irrelevantes. Isso ocorre porque, à medida que os dados aumentam, o desempenho da consulta pode diminuir com o tempo, principalmente devido ao grande volume de dados a serem verificados quando uma consulta é executada. O Amazon S3 e outros serviços oferecem orientação para a migração do ciclo de vida dos dados, que mostra estratégias para mover dados para diferentes locais de armazenamento à medida que esfriam. Também há um benefício de custo de armazenamento em fazer isso.

Além da migração de dados, você pode usar outras estratégias, como remover dados de origem que não são relevantes para as consultas que você está executando. Isso pode dar algum trabalho, pois pode significar alterar seu esquema de dados de origem. Mas seu resultado positivo é reduzir o volume de dados e resultar em consultas mais rápidas. Para obter mais informações, consulte Gerenciar o ciclo de vida de objetos.

Trino

Considere o seguinte ao executar o Trino no Amazon EMR.

Propriedades de implantação do Trino não configuráveis

A tabela a seguir mostra as diferentes opções de configuração para `properties` arquivos Trino.

Arquivo	Configurável
`log.properties`	Trino: configurável nas versões 6.1.0 e posteriores do Amazon EMR. Use a classific ação de configuração `prestosql-log` ou `trino-log` .
`config.properties`	Trino: configurável nas versões 6.1.0 e posteriores do Amazon EMR. Use a classific ação de configuração `prestosql-config` ou `trino-config` .
`hive.properties`	Trino: configurável nas versões 6.1.0 e posteriores do Amazon EMR. Use a classific ação de configuração `prestosql-connector-hive` ou `trino-connector-hive` .
`node.properties`	Trino: configurável nas versões 6.1.0 e posteriores do Amazon EMR. Use a classific ação de configuração `prestosql-node` ou `trino-node` .
`jvm.config`	Não configurável.

Considerações adicionais

- Para o Trino nas versões 6.1.0 e posteriores do EMR, o Amazon EMR configura automaticamente uma chave secreta compartilhada para comunicação interna segura entre os nós do cluster. Você não precisa fazer qualquer configuração adicional para habilitar esse atributo de segurança e

pode substituir a configuração com sua própria chave secreta. Para obter informações sobre a autenticação interna do Trino, consulte a documentação do Trino 353: Proteger a comunicação interna.

Histórico de versões do Trino

As seções de notas de lançamento detalham as mudanças e atualizações de uma versão específica do Trino no Amazon EMR.

Trino por versão

- Amazon EMR 7.6.0: notas de versão do Trino
- Amazon EMR 7.3.0: notas de versão do Trino
- Amazon EMR 6.6.0: notas de versão do Trino

Amazon EMR 7.6.0: notas de versão do Trino

Amazon EMR 7.6.6.6.6.0: recursos do Trino

- Para dar suporte a consultas de longa execução, o Trino agora inclui um mecanismo de execução tolerante a falhas. A execução tolerante a falhas atenua as falhas nas consultas ao tentar novamente as consultas com falha ou as tarefas dos seus componentes.

Amazon EMR 7.6.6.0: alterações no Trino

Amazon EMR 7.6.6.0: alterações no Trino

Tipo	Descrição
Upgrade	Atualização do Trino para 457

Amazon EMR 7.3.0: notas de versão do Trino

Amazon EMR 7.3.0: alterações no Trino

- Essa versão atualiza o Trino da versão 436 para a 442.

- Essa versão redireciona as consultas do Hudi para o novo corretor do Hudi. O antigo conector do Hive não consegue mais ler tabelas do Hudi. Observação
- Essa versão remove o módulo Rubix do Amazon EMR porque ele já está obsoleto do código aberto.
- Essa versão remove o modo legado na propriedade `hive.security`. O padrão agora é `allow-all`.

Amazon EMR 6.6.0: notas de versão do Trino

Amazon EMR 6.6.6.0: recursos do Trino

- Para dar suporte a consultas de longa execução, o Trino agora inclui um mecanismo de execução tolerante a falhas. A execução tolerante a falhas atenua as falhas nas consultas ao tentar novamente as consultas com falha ou as tarefas dos seus componentes.

Amazon EMR 6.9.0: alterações no Trino

Amazon EMR 6.9.0: alterações no Trino

Tipo	Descrição
Upgrade	Atualização do Trino para 398
Upgrade	Suporte para Hadoop 3.3.3
Recurso	Suporte ao Tardigrade: adicione suporte para spooling de troca no HDFS e no Amazon S3.
Correção de bugs	Quando o Iceberg do Trino é usado e o catálogo do Glue está habilitado, evite adicionar o uri do metastore em `iceberg.properties`.

Amazon EMR 6.6.0: problemas conhecidos do Trino

- Para a versão 6.9.0 do Amazon EMR, o Trino não funciona em clusters habilitados para o Apache Ranger. Se você precisar usar o Trino com o Ranger, entre em contato com o Suporte.

Apache Zeppelin

Use o Apache Zeppelin como um notebook para exploração de dados interativos. Para obter mais informações sobre o Zeppelin, consulte https://zeppelin.apache.org/. O Zeppelin é incluído nas versões 5.0.0 e posteriores do Amazon EMR. As versões anteriores incluem o Zeppelin como uma aplicação em sandbox. Para obter mais informações, consulte Versões de lançamento 4.x do Amazon EMR.

Para acessar a interface da web Zeppelin, configure um túnel SSH para o nó principal e uma conexão de proxy. Para obter mais informações, consulte Visualizar interfaces da Web hospedadas em clusters do EMR.

A tabela a seguir lista a versão do Zeppelin incluída na versão mais recente da série 7.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Zeppelin.

Para obter a versão dos componentes instalados com o Zeppelin nesse lançamento, consulte Release 7.9.0 Component Versions.

Informações de versão do Zeppelin para o emr-7.9.0

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-7.9.0	Zeppelin 0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

A tabela a seguir lista a versão do Zeppelin incluída na versão mais recente da série 6.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Zeppelin.

Para obter a versão dos componentes instalados com o Zeppelin nessa versão, consulte Release 6.15.0 Component Versions.

Informações de versão do Zeppelin para o emr-6.15.0

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-6.15.0	Zeppelin 0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

A tabela a seguir lista a versão do Zeppelin incluída na versão mais recente da série 5.x do Amazon EMR, além dos componentes que o Amazon EMR instala com o Zeppelin.

Para obter a versão dos componentes instalados com o Zeppelin nesse lançamento, consulte Release 5.36.2 Component Versions.

Informações de versão do Zeppelin para o emr-5.36.2

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.36.2	Zeppelin 0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
		client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Tópicos

- Considerações ao usar o Zeppelin no Amazon EMR
- Histórico de versões do Zeppelin

Considerações ao usar o Zeppelin no Amazon EMR

- Conecte-se ao Zeppelin usando o mesmo método de encapsulamento SSH para conectar-se a outros servidores web no nó principal. O servidor do Zeppelin está disponível na porta 8890.
- O Zeppelin nas versões 5.0.0 e posteriores do Amazon EMR é compatível com a autenticação Shiro.
- O Zeppelin nas versões 5.8.0 e posteriores do Amazon EMR no Amazon EMR versão 5.8.0 e versões posteriores é compatível com o uso do Catálogo de Dados do AWS Glue como o metastore do Spark SQL. Para obter mais informações, consulte Using AWS Glue Data Catalog as the metastore for Spark SQL.
- O Zeppelin não usa algumas das configurações definidas no arquivo de configuração `spark-defaults.conf` do seu cluster, embora ele instrua o YARN para alocar executores dinamicamente se você definir `spark.dynamicAllocation.enabled` como `true`. Você deve definir as configurações do executor, como memória e núcleos, usando a guia Interpreter (Intérprete) do Zeppelin e reiniciar o intérprete para que eles sejam usados.

- O Amazon EMR 6.10.0 e versões posteriores oferecem suporte à integração do Apache Zeppelin com o Apache Flink. Consulte Usar trabalhos do Flink pelo Zeppelin no Amazon EMR para obter mais informações.
- O Zeppelin no Amazon EMR não é compatível com o intérprete do SparkR.

Histórico de versões do Zeppelin

A tabela a seguir lista a versão do Zeppelin incluída em cada versão do Amazon EMR, além dos componentes instalados com a aplicação. Para obter as versões dos componentes em cada versão, consulte a seção Versão de componentes da sua versão em Versões 7.x do Amazon EMR, Versões de lançamento 6.x do Amazon EMR ou Versões de lançamento 5.x do Amazon EMR.

Informações da versão do Zeppelin

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-7.9.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-7.8.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodema

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
		nager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-7.7.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-7.6.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-7.5.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-7.4.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-7.3.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-7.2.0	0.10.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.36.2	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-7.1.0	0.10.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-7.0.0	0.10.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-6.15.0	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.14.0	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-6.13.0	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.12.0	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-6.11.1	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.11.0	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-6.10.1	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.10.0	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-6.9.1	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.9.0	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-6.8.1	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.8.0	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-6.7.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.36.1	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.36.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.6.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.35.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.5.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-6.4.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.3.1	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-6.3.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.2.1	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-6.2.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.1.1	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-6.1.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.0.1	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-6.0.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.34.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.33.1	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.33.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.32.1	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.32.0	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.31.1	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.31.0	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.30.2	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.30.1	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.30.0	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.29.0	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.28.1	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.28.0	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.27.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.27.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.26.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.25.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.24.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.24.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.23.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.23.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.22.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.21.2	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.21.1	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.21.0	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.20.1	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.20.0	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.19.1	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.19.0	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.18.1	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.18.0	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.17.2	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.17.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.17.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.16.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.16.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.15.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.15.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.14.2	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.14.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.14.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.13.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.13.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.12.3	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.12.2	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.12.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.12.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.11.4	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.11.3	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.11.2	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.11.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.11.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.10.1	0.7.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.10.0	0.7.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.9.1	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.9.0	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.8.3	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.8.2	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.8.1	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.8.0	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.7.1	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.7.0	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.6.1	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.6.0	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.5.4	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.5.3	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.5.2	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.5.1	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.5.0	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.4.1	0.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.4.0	0.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.3.2	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.3.1	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.3.0	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.2.3	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.2.2	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.2.1	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.2.0	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.1.1	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.1.0	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.0.3	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.0.2	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.0.1	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rótulo de versão do Amazon EMR	Versão do Zeppelin	Componentes instalados com o Zeppelin
emr-5.0.0	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Apache ZooKeeper

O Apache ZooKeeper é um serviço centralizado para manter informações de configuração, nomear, fornecer sincronização distribuída e oferecer serviços de grupo. Para obter mais informações sobre ZooKeeper, consulte http://zookeeper.apache.org/.

A tabela a seguir lista a versão ZooKeeper incluída na versão mais recente da série 7.x do Amazon EMR, além dos componentes com os quais o Amazon EMR instala. ZooKeeper

Para obter a versão dos componentes instalados ZooKeeper nessa versão, consulte Release 7.9.0 Component Versions.

ZooKeeper informações de versão para o emr-7.9.0

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-7.9.0	ZooKeeper 3.9.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

A tabela a seguir lista a versão ZooKeeper incluída na versão mais recente da série 6.x do Amazon EMR, além dos componentes com os quais o Amazon EMR instala. ZooKeeper

Para obter a versão dos componentes instalados ZooKeeper nessa versão, consulte Release 6.15.0 Component Versions.

ZooKeeper informações de versão para o emr-6.15.0

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-6.15.0	ZooKeeper 3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

A tabela a seguir lista a versão ZooKeeper incluída na versão mais recente da série 5.x do Amazon EMR, além dos componentes com os quais o Amazon EMR instala. ZooKeeper

Para obter a versão dos componentes instalados ZooKeeper nessa versão, consulte Release 5.36.2 Component Versions.

ZooKeeper informações de versão para o emr-5.36.2

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.36.2	ZooKeeper 3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Tópicos

- ZooKeeper histórico de versões
- ZooKeeper notas da versão

ZooKeeper histórico de versões

A tabela a seguir lista a versão ZooKeeper incluída em cada versão do Amazon EMR, além dos componentes instalados com a aplicação. Para obter as versões dos componentes em cada versão, consulte a seção Versão de componentes da sua versão em Versões 7.x do Amazon EMR, Versões de lançamento 6.x do Amazon EMR ou Versões de lançamento 5.x do Amazon EMR.

ZooKeeper informações sobre a versão

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-7.9.0	3.9.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-7.8.0	3.9.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-7.7.0	3.9.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-7.6.0	3.9.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-7.5.0	3.9.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-7.4.0	3.9.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-7.3.0	3.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-7.2.0	3.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.36.2	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-7.1.0	3.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-7.0.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-6.15.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.14.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.13.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-6.12.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.11.1	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.11.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-6.10.1	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.10.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.9.1	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-6.9.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.8.1	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.8.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-6.7.0	3.5.7	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.36.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.36.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-6.6.0	3.5.7	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.35.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.5.0	3.5.7	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-6.4.0	3.5.7	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.3.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.3.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-6.2.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.2.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.1.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-6.1.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.0.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.0.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.34.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.33.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.33.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.32.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.32.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.31.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.31.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.30.2	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.30.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.30.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.29.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.28.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.28.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.27.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.27.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.26.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.25.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.24.1	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.24.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.23.1	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.23.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.22.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.21.2	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.21.1	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.21.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.20.1	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.20.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.19.1	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.19.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.18.1	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.18.0	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.17.2	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.17.1	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.17.0	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.16.1	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.16.0	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.15.1	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.15.0	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.14.2	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.14.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.14.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.13.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.13.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.12.3	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.12.2	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.12.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.12.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.11.4	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.11.3	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.11.2	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.11.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.11.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.10.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.10.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.9.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.9.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.8.3	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.8.2	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.8.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.8.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.7.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.7.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.6.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.6.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.5.4	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.5.3	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.5.2	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.5.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.5.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.4.1	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.4.0	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.3.2	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.3.1	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.3.0	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.2.3	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.2.2	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.2.1	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.2.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.1.1	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.1.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.0.3	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Rótulo de versão do Amazon EMR	ZooKeeper Versão	Componentes instalados com ZooKeeper
emr-5.0.2	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.0.1	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.0.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

ZooKeeper notas da versão

Consulte as seções a seguir para ver as notas de lançamento completas.

Amazon EMR 7.8.0 - Notas da versão ZooKeeper

Amazon EMR 7.8.0 - Alterações ZooKeeper

Atualização: o Zookeeper foi atualizado para a versão 3.9.3

Amazon EMR 7.5.0: notas da versão ZooKeeper

Amazon EMR 7.5.0: alterações no Amazon EMR 7.5.0 - Alterações ZooKeeper

- A partir do EMR-7.5.0, o Zookeeper é configurado para Java 17 por padrão em tempo de execução. Para usar uma versão diferente para o Java Runtime, substitua as configurações da JVM no `zookeeper-server` arquivo e defina `JAVA_HOME` para a versão desejada.

Amazon EMR 7.4.0: notas da versão ZooKeeper

Tipo	Descrição
Upgrade	A versão do Zookeeper foi atualizada para 3.9.2
Melhoria	ZOOKEEPER-4778: Corrija o jetty, o netty e o logback para remover vulnerabilidades de alta gravidade.
Melhoria	ZOOKEEPER-4799: corrige CVE-2024-23944.
Correção de bugs	O suporte ZOOKEEPER-4415: TLSv1 .3 foi adicionado.

Amazon EMR 7.4.0: recursos ZooKeeper

- Clusters com configuração de segurança de criptografia em trânsito fornecida terão o TLS habilitado por padrão para comunicação de quórum e Zookeeper Admin Server. Além disso, a porta segura TLS é exposta em 2281 para comunicação do cliente com o Zookeeper.

Conectores e utilitários

O Amazon EMR fornece vários conectores e utilitários para acessar outros AWS serviços como fontes de dados. Você pode geralmente acessar os dados nesses serviços dentro de um programa. Por exemplo, você pode especificar um stream do Kinesis em uma consulta, script do Pig ou MapReduce aplicativo do Hive e, em seguida, operar com esses dados.

Tópicos

- Exportar, importar, consultar e unir tabelas no DynamoDB usando o Amazon EMR
- Kinesis
- S3 DistCp (s3-dist-cp)
- Limpando após falhas em trabalhos do S3 DistCp

Exportar, importar, consultar e unir tabelas no DynamoDB usando o Amazon EMR

> **Note**
>
> O Amazon EMR-DynamoDB Connector é de código aberto em. GitHub Para obter mais informações, consulte https://github.com/awslabs/emr-dynamodb-connector.

O DynamoDB é um serviço de banco de dados NoSQL totalmente gerenciado que proporciona uma performance rápida e previsível com escalabilidade contínua. Os desenvolvedores podem criar uma tabela de banco de dados e ampliar seu tráfego de solicitação ou armazenamento sem limites. O DynamoDB distribui automaticamente dados e tráfego para a tabela através de um número suficiente de servidores, a fim de controlar a capacidade de solicitação especificada pelo cliente e a quantidade de dados armazenados, enquanto mantém um desempenho consistente e rápido. Usando o Amazon EMR e o Hive, você pode processar grandes quantidades de dados com rapidez e eficiência, como os dados armazenados no DynamoDB. Para obter mais informações sobre o DynamoDB, consulte o Guia do desenvolvedor do Amazon DynamoDB.

Apache Hive é uma camada de software que você pode usar para consultar clusters do map-reduce usando uma linguagem simplificada semelhante ao SQL chamada HiveQL. Ele é executado sobre a arquitetura do Hadoop. Para obter mais informações sobre o Hive e o HiveQL, acesse o documento

HiveQL Language Manual. Para obter mais informações sobre o Hive e o Amazon EMR, consulte Apache Hive.

Você pode usar o Amazon EMR com uma versão personalizada do Hive que inclua conectividade com o DynamoDB para executar operações em dados armazenados no DynamoDB:

- Carregar dados do DynamoDB no Sistema de Arquivos Distribuído do Hadoop (HDFS) e usá-los como entrada em um cluster do Amazon EMR.
- Consultar dados dinâmicos do DynamoDB usando instruções semelhantes a SQL (HiveQL).
- Unir dados armazenados no DynamoDB e exportar ou consultar com base nos dados unidos.
- Exportar dados armazenados no DynamoDB para o Amazon S3.
- Importar dados armazenados no Amazon S3 para o DynamoDB.

> **Note**
>
> O conector Amazon EMR-DynamoDB não é compatível com os clusters configurados para usar a autenticação Kerberos.

Para executar cada uma das seguintes tarefas, será necessário iniciar um cluster do Amazon EMR, especificar a localização dos dados no DynamoDB e emitir comandos do Hive para manipular os dados no DynamoDB.

Há várias maneiras de iniciar um cluster do Amazon EMR: você pode usar o console do Amazon EMR, a interface de linha de comando (CLI) ou pode programar seu cluster usando um SDK AWS ou a API do Amazon EMR. Você também pode escolher se deseja executar um cluster do Hive interativamente ou a partir de um script. Nesta seção, vamos mostrar a você como iniciar um cluster interativo do Hive pelo console do Amazon EMR e pela CLI.

Usar o Hive, interativamente, é uma ótima maneira de testar o desempenho das consultas e ajustar seu aplicativo. Depois que você tiver estabelecido um conjunto de comandos do Hive que serão executados de forma regular, considere criar um script do Hive que o Amazon EMR possa executar para você.

> **Warning**
>
> As operações de leitura ou de gravação do Amazon EMR em uma tabela do DynamoDB contam em relação ao throughput provisionado estabelecido, o que pode aumentar a

frequência de exceções de throughput provisionado. Para grandes solicitações, o Amazon EMR implementa novas tentativas com recuo exponencial para gerenciar a carga de solicitações na tabela do DynamoDB. Executar trabalhos do Amazon EMR simultaneamente com outro tráfego pode fazer com que você exceda o nível de throughput provisionado alocado. Você pode monitorar isso verificando a ThrottleRequestsmétrica na Amazon CloudWatch. Se a carga de solicitações for muito alta, você pode reiniciar o cluster e definir Configuração de porcentagem de leitura ou Configuração de porcentagem de gravação como um valor menor para limitar as operações do Amazon EMR. Para obter mais informações sobre as configurações de throughput do DynamoDB, consulte Provisioned throughput.
Se uma tabela estiver configurada para o Modo sob demanda, será necessário alterá-la novamente para o modo provisionado antes de executar uma operação de exportação ou importação. Os pipelines precisam de uma taxa de produtividade para calcular os recursos a serem usados em um Dynamo. DBtable O modo sob demanda remove o throughput provisionado. Para provisionar a capacidade de transferência, você pode usar as métricas do Amazon CloudWatch Events para avaliar a taxa de transferência agregada usada por uma tabela.

Tópicos

- Configurar uma tabela do Hive para executar comandos do Hive
- Exemplos de comandos do Hive para exportar, importar e consultar dados no DynamoDB
- Otimizar a performance de operações do Amazon EMR no DynamoDB

Configurar uma tabela do Hive para executar comandos do Hive

O Apache Hive é uma aplicação de data warehouse que você pode usar para consultar dados contidos em clusters do Amazon EMR usando uma linguagem semelhante ao SQL. Para obter mais informações sobre o Hive, consulte http://hive.apache.org/.

O procedimento a seguir pressupõe que você já tenha criado um cluster e especificado um par de EC2 chaves da Amazon. Para saber como começar a criar clusters, consulte Getting started with Amazon EMR no Guia de gerenciamento do Amazon EMR.

Configurar o Hive para usar MapReduce

Quando você usar o Hive no Amazon EMR para consultar tabelas do DynamoDB, poderão ocorrer erros se o Hive usar o mecanismo de execução padrão, o Tez. Por esse motivo, ao criar um cluster

com o Hive que se integre ao DynamoDB conforme descrito nesta seção, recomendamos que você use uma classificação de configuração que defina o Hive a ser usado. MapReduce Para obter mais informações, consulte Configurar aplicações.

O trecho a seguir mostra a classificação de configuração e a propriedade a serem usadas para definir MapReduce como mecanismo de execução do Hive:

```
[
                {
                    "Classification": "hive-site",
                    "Properties": {
                        "hive.execution.engine": "mr"
                    }
                }
            ]
```

Para executar comandos do Hive interativamente

1. Conecte-se ao nó principal. Para obter mais informações, consulte Conectar-se ao nó principal usando SSH no Guia de gerenciamento do Amazon EMR.
2. No prompt de comando do nó principal atual, digite `hive`.

   Você verá um prompt do Hive: `hive>`
3. Insira um comando do Hive que mapeie uma tabela na aplicação Hive para os dados no DynamoDB. Essa tabela serve de referência para os dados armazenados no Amazon DynamoDB; os dados não são armazenados localmente no Hive, e qualquer consulta usando essa tabela é executada com base nos dados dinâmicos no DynamoDB, consumindo a capacidade de leitura ou gravação da tabela sempre que um comando é executado. Se você pretende executar vários comandos do Hive no mesmo conjunto de dados, considere exportá-lo primeiro.

   O exemplo a seguir mostra a sintaxe para mapear uma tabela do Hive para uma tabela do DynamoDB.

   ```
   CREATE EXTERNAL TABLE hive_tablename
    (hive_column1_name column1_datatype, hive_column2_name column2_datatype...)
   STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
   TBLPROPERTIES ("dynamodb.table.name" = "dynamodb_tablename",
   ```

```
"dynamodb.column.mapping" =
 "hive_column1_name:dynamodb_attribute1_name,hive_column2_name:dynamodb_attribute2_name..."
```

Ao criar uma tabela no Hive pelo DynamoDB, você deve criá-la como uma tabela externa usando a palavra-chave EXTERNAL. A diferença entre tabelas externas e internas é que os dados nas tabelas internas são excluídos quando uma tabela interna é descartada. Este não é o comportamento desejado no caso de uma conexão com o Amazon DynamoDB e, portanto, somente tabelas externas são compatíveis.

Por exemplo, o seguinte comando do Hive cria uma tabela chamada hivetable1 no Hive que referencia a tabela do DynamoDB chamada dynamodbtable1. A tabela dynamodbtable1 do DynamoDB tem um esquema de chave primária. hash-and-range O elemento da chave de hash é `name` (do tipo string), o elemento de chave de intervalo é `year` (do tipo numérico), e cada item possui um valor de atributo para `holidays` (do tipo conjunto de strings).

```
CREATE EXTERNAL TABLE hivetable1 (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");
```

A linha 1 usa a instrução HiveQL `CREATE EXTERNAL TABLE`. Para hivetable1, você precisa estabelecer uma coluna para cada par de nome-valor de atributo na tabela do DynamoDB e fornecer o tipo de dados. Esses valores não diferenciam maiúsculas de minúsculas, e você pode dar qualquer nome (exceto palavras reservadas) para as colunas.

A linha 2 usa a instrução `STORED BY`. O valor de `STORED BY` é o nome da classe que manipula a conexão entre o Hive e o DynamoDB. Ele deve ser definido como `'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'`.

A linha 3 usa a instrução `TBLPROPERTIES` para associar “hivetable1” com a tabela e o esquema corretos no DynamoDB. Forneça a `TBLPROPERTIES` valores para os parâmetros `dynamodb.table.name` e `dynamodb.column.mapping`. Esses valores diferenciam maiúsculas de minúsculas.

> **Note**
>
> Todos os nomes de atributo do DynamoDB para a tabela devem ter colunas correspondentes na tabela do Hive. Dependendo da sua versão do Amazon EMR, os seguintes cenários ocorrem se o one-to-one mapeamento não existir:
>
> - No Amazon EMR versão 5.27.0 e posterior, o conector tem validações que garantem um mapeamento one-to-one entre nomes de atributos e colunas do DynamoDB na tabela do Hive. Ocorrerá um erro se o one-to-one mapeamento não existir.
> - Na versão 5.26.0 do Amazon EMR e anteriores, a tabela do Hive não contém o par de nome-valor do DynamoDB. Se você não mapear os atributos de chave primária do DynamoDB, o Hive gerará um erro. Se você não mapear um atributo de chave não primária, nenhum erro será gerado, mas os dados não serão vistos na tabela do Hive. Se os tipos de dados não corresponderem, o valor será nulo.

Em seguida, você poderá começar a executar operações do Hive em hivetable1. As consultas executadas em hivetable1 são internamente executadas na tabela do DynamoDB dynamodbtable1 da sua conta do DynamoDB, consumindo unidades de leitura ou gravação com cada execução.

Quando você executa consultas do Hive em uma tabela do DynamoDB, precisa certificar-se de ter provisionado uma quantidade suficiente de unidades de capacidade de leitura.

Por exemplo, suponha que você tenha provisionado 100 unidades de capacidade de leitura para a sua tabela do DynamoDB. Isso permitirá que você realize 100 leituras, ou 409.600 bytes, por segundo. Se essa tabela contiver 20 GB de dados (21.474.836.480 bytes) e sua consulta do Hive realizar uma verificação de tabela completa, você poderá estimar quanto tempo a consulta demorará para ser executada:

21.474.836.480/409.600 = 52.429 segundos = 14,56 horas

A única maneira de diminuir o tempo necessário seria ajustar as unidades de capacidade de leitura na tabela DynamoDB de origem. Não adianta adicionar mais nós do Amazon EMR.

Na saída do Hive, a porcentagem de conclusão é atualizada quando um ou mais processos de mapeador são finalizados. Para uma tabela grande do DynamoDB com uma configuração baixa de capacidade de leitura provisionada, a saída de percentual de conclusão pode não ser atualizada por um longo tempo. No caso acima, o trabalho parecerá estar 0% concluído por várias horas. Para ver

um status mais detalhado sobre o andamento do trabalho, acesse o console do Amazon EMR. Você poderá visualizar o status de tarefas de mapeador individuais e as estatísticas de leituras de dados. Você também pode fazer logon na interface do Hadoop no nó principal e visualizar as estatísticas do Hadoop. Isso mostrará o status de tarefas map individuais e algumas estatísticas de leitura de dados. Para obter mais informações, consulte os tópicos a seguir.

- Interfaces Web hospedadas no nó principal
- Visualizar as interfaces Web do Hadoop

Para obter mais informações sobre instruções HiveQL de exemplo para realizar tarefas, como exportar ou importar dados do DynamoDB e unir tabelas, consulte Exemplos de comandos do Hive para exportar, importar e consultar dados no DynamoDB.

Para cancelar uma solicitação do Hive

Quando você executa uma consulta do Hive, a resposta inicial do servidor inclui o comando para cancelar a solicitação. Para cancelar a solicitação a qualquer momento no processo, use Kill Command (Comando Kill) na resposta do servidor.

1. Insira `Ctrl+C` para sair do cliente de linha de comando.
2. No prompt do shell, insira o Kill Command (Comando Kill) da resposta do servidor inicial à sua solicitação.

   Como alternativa, você pode executar o seguinte comando na linha de comando do nó principal para eliminar o trabalho do Hadoop, onde *job-id* é o identificador do trabalho do Hadoop e pode ser recuperado da interface do usuário do Hadoop.

   ```
   hadoop job -kill job-id
   ```

Tipos de dados para o Hive e o DynamoDB

A tabela a seguir mostra os tipos de dados do Hive disponíveis, o tipo do DynamoDB padrão ao qual eles correspondem e os tipos do DynamoDB alternativos para os quais eles também podem mapear.

Tipo do Hive	Tipo do DynamoDB padrão	Tipos alternativos do DynamoDB
string	string (S)	
bigint ou duplo	número (N)	
binary	binário (B)	
boolean	booliano (BOOL)	
array	lista (L)	conjunto de números (NS), conjunto de strings (SS) ou conjunto de binários (BS)
mapa<string, string>	item	mapa (M)
mapa<string,?>	mapa (M)	
	nulo (NULL)	

Se você deseja gravar seus dados do Hive como um tipo alternativo correspondente do DynamoDB, ou se os dados do DynamoDB contiverem valores de atributo de um tipo alternativo do DynamoDB, você poderá especificar a coluna e o tipo do DynamoDB usando o parâmetro `dynamodb.type.mapping`. O exemplo a seguir mostra a sintaxe para especificar um tipo alternativo de mapeamento.

```
CREATE EXTERNAL TABLE hive_tablename (hive_column1_name column1_datatype,
 hive_column2_name column2_datatype...)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodb_tablename",
"dynamodb.column.mapping" =
 "hive_column1_name:dynamodb_attribute1_name,hive_column2_name:dynamodb_attribute2_name...",
"dynamodb.type.mapping" = "hive_column1_name:dynamodb_attribute1_datatype");
```

O parâmetro de mapeamento de tipo é opcional e só tem que ser especificado para as colunas que usam tipos alternativos.

Por exemplo, o comando do Hive a seguir cria uma tabela chamada `hivetable2` que referencia tabela `dynamodbtable2` do DynamoDB. Ela é semelhante à `hivetable1`, exceto pelo fato de mapear a coluna `col3` para o tipo de conjunto de strings (SS).

```
CREATE EXTERNAL TABLE hivetable2 (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable2",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays",
"dynamodb.type.mapping" = "col3:SS");
```

No Hive, `hivetable1` e `hivetable2` são idênticas. No entanto, quando os dados dessas tabelas forem gravados em suas tabelas correspondentes do DynamoDB, a `dynamodbtable1` conterá listas, enquanto a `dynamodbtable2` conterá conjuntos de strings.

Se você quiser gravar valores `null` do Hive como atributos do tipo `null` do DynamoDB, poderá fazer isso usando o parâmetro `dynamodb.null.serialization`. O exemplo a seguir mostra a sintaxe para especificar a serialização `null`.

```
CREATE EXTERNAL TABLE hive_tablename (hive_column1_name column1_datatype,
 hive_column2_name column2_datatype...)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodb_tablename",
"dynamodb.column.mapping" =
 "hive_column1_name:dynamodb_attribute1_name,hive_column2_name:dynamodb_attribute2_name...",
"dynamodb.null.serialization" = "true");
```

O parâmetro de serialização nula é opcional e será definido como `false` se não especificado. Os atributos `null` do DynamoDB são lidos como valores `null` no Hive, independentemente da configuração do parâmetro. As coleções do Hive com valores `null` só poderão ser gravadas no DynamoDB se o parâmetro de serialização nula for especificado como `true`. Caso contrário, ocorrerá um erro do Hive.

O tipo bigint no Hive equivale ao tipo longo no Java, enquanto o tipo duplo no Hive equivale ao tipo duplo no Java em termos de precisão. Isso significa que, se você tiver dados numéricos armazenados no DynamoDB cuja precisão seja maior que a disponível nos tipos de dados do Hive, usar o Hive para exportar, importar ou referenciar os dados do DynamoDB pode resultar na perda de precisão ou em uma falha da consulta do Hive.

As exportações do tipo binário do DynamoDB para o Amazon Simple Storage Service (Amazon S3) ou o HDFS são armazenadas como uma string codificada em Base64. Se você estiver importando

dados do Amazon S3 ou do HDFS para o tipo binário do DynamoDB, eles deverão ser codificados como uma string Base64.

Opções do Hive

Você pode definir as seguintes opções do Hive para gerenciar a transferência de dados do Amazon DynamoDB. Essas opções só persistem na sessão atual do Hive. Se você fechar o prompt de comando do Hive e reabri-lo mais tarde no cluster, essas configurações retornarão aos valores padrão.

Opções do Hive	Descrição
`dynamodb.throughput.read.percent`	Defina a taxa de operações de leitura para manter sua taxa de throughput provisionado do DynamoDB no intervalo alocado para sua tabela. O valor está entre `0.1` e `1.5`, inclusive. O valor de 0,5 é a taxa de leitura padrão, que significa que o Hive tentará consumir metade dos recursos de throughput provisionado de leitura na tabela. Aumentar esse valor acima de 0,5 aumenta a taxa de solicitações de leitura. Diminuí-lo abaixo de 0,5 diminui a taxa de solicitações de leitura. Essa taxa de leituras é aproximada. A taxa de leituras real dependerá de fatores como, por exemplo, a existência de distribuição uniforme de chaves no DynamoDB. Se você perceber que o seu throughput provisionado está sendo frequentemente excedido pela operação do Hive, ou se o tráfego de leituras dinâmicas estiver muito limitado, reduza esse valor abaixo de `0.5`. Se você tiver capacidade suficiente e quiser uma operação do Hive mais rápida, defina esse valor acima de `0.5`. Também é possível configurar o excesso de assinaturas, definindo o valor como até 1.5 se você acredita que existem operações de entrada/saída não utilizadas disponíveis.

Opções do Hive	Descrição
`dynamodb.throughput.write.percent`	Defina a taxa de operações de gravação para manter sua taxa de throughput provisionado do DynamoDB no intervalo alocado para sua tabela. O valor está entre `0.1` e `1.5`, inclusive. O valor de 0,5 é a taxa de gravação padrão, que significa que o Hive tentará consumir metade dos recursos de throughput provisionado de gravação na tabela. Aumentar esse valor acima de 0,5 aumenta a taxa de solicitações de gravação. Diminuí-lo abaixo de 0,5 diminui a taxa de solicitações de gravação. Essa taxa de gravações é aproximada. A taxa de gravações real dependerá de fatores como se há uma distribui ção uniforme de chaves no DynamoDB Se você perceber que o seu throughput provisionado está sendo frequentemente excedido pela operação do Hive, ou se o tráfego de gravações dinâmicas estiver muito limitado, reduza esse valor abaixo de `0.5`. Se você tiver capacidade suficiente e quiser uma operação do Hive mais rápida, defina esse valor acima de `0.5`. Também é possível configurar o excesso de assinaturas, definindo o valor como até 1.5 se você acredita que existem operações de entrada/saída não utilizadas disponíveis ou se este for o upload de dados inicial na tabela e ainda não houver tráfego dinâmico.
`dynamodb.endpoint`	Especifique o endpoint para o serviço do DynamoDB. Para obter mais informações sobre os endpoints do DynamoDB disponíveis, consulte Regions and endpoints.
`dynamodb.max.map.tasks`	Especifique o número máximo de tarefas map ao ler dados do DynamoDB. Esse valor deve ser igual ou maior que 1.

Opções do Hive	Descrição
`dynamodb.retry.duration`	Especifique o número de minutos a serem usados como a duração do tempo limite para tentar novamente os comandos do Hive. Esse valor deve ser um número inteiro igual a ou maior que 0. A duração do tempo limite padrão é de dois minutos.

Essas opções são definidas usando o comando `SET`, como mostra o exemplo a seguir.

```
SET dynamodb.throughput.read.percent=1.0;

INSERT OVERWRITE TABLE s3_export SELECT *
FROM hiveTableName;
```

Exemplos de comandos do Hive para exportar, importar e consultar dados no DynamoDB

Os exemplos a seguir usam comandos do Hive para realizar operações, como exportar dados para o Amazon S3 ou HDFS, importar dados para o DynamoDB, unir tabelas, consultar tabelas e muito mais.

As operações em uma tabela do Hive referenciam dados armazenados no DynamoDB. Os comandos do Hive estão sujeitos às configurações de throughput provisionado da tabela do DynamoDB, e os dados recuperados incluem os dados gravados na tabela do DynamoDB na ocasião em que a solicitação de operação do Hive é processada pelo DynamoDB. Se o processo de recuperação de dados demorar muito, alguns dados retornados pelo comando do Hive podem ter sido atualizados no DynamoDB desde que o comando do Hive foi iniciado.

Os comandos do Hive `DROP TABLE` e `CREATE TABLE` atuam apenas nas tabelas locais do Hive e não criam nem descartam tabelas no DynamoDB. Se a sua consulta do Hive fizer referência a uma tabela no DynamoDB, essa tabela já deverá existir antes da execução da consulta. Para obter mais informações sobre como criar e excluir tabelas no Amazon DynamoDB, consulte Working with tables in DynamoDB no Guia do desenvolvedor do Amazon DynamoDB.

> **Note**
>
> Quando você mapear uma tabela do Hive para um local no Amazon S3, não a mapeie para o caminho raiz do bucket, s3://amzn-s3-demo-bucket, pois isso poderá causar erros quando o Hive gravar os dados no Amazon S3. Em vez disso, mapeie a tabela para um subcaminho do bucket, s3://amzn-s3-demo-bucket/mypath.

Exportar dados do DynamoDB

Você pode usar o Hive para exportar dados do DynamoDB.

Exportar uma tabela do DynamoDB para um bucket do Amazon S3

- Crie uma tabela do Hive que faça referência aos dados armazenados no DynamoDB. Em seguida, você poderá chamar o comando INSERT OVERWRITE para gravar os dados em um diretório externo. No exemplo a seguir, *s3://amzn-s3-demo-bucket/path/subpath/* é um caminho válido no Amazon S3. Ajuste as colunas e os tipos de dados no comando CREATE para que eles correspondam aos valores no DynamoDB. Você pode usar isso para criar um arquivamento dos seus dados do DynamoDB no Amazon S3.

```
CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");


INSERT OVERWRITE DIRECTORY 's3://amzn-s3-demo-bucket/path/subpath/' SELECT *
FROM hiveTableName;
```

Exportar uma tabela do DynamoDB para um bucket do Amazon S3 usando formatação

- Crie uma tabela externa que referencie um local no Amazon S3. Isso é mostrado abaixo como s3_export. Durante a chamada CREATE, especifique a formatação de linhas para a tabela. Em seguida, quando você usar INSERT OVERWRITE para exportar dados do DynamoDB para o s3_export, os dados serão gravados no formato especificado. No exemplo a seguir, os dados são gravados como valores separados por vírgula (CSV).

```
CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");


CREATE EXTERNAL TABLE s3_export(a_col string, b_col bigint, c_col array<string>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';

INSERT OVERWRITE TABLE s3_export SELECT *
FROM hiveTableName;
```

Exportar uma tabela do DynamoDB para um bucket do Amazon S3 sem especificar um mapeamento de colunas

- Crie uma tabela do Hive que faça referência aos dados armazenados no DynamoDB. Isso é semelhante ao exemplo anterior, com a diferença de que você não está especificando um mapeamento de colunas. A tabela deve ter exatamente uma coluna do tipo `map<string, string>`. Se você criar uma tabela `EXTERNAL` no Amazon S3, poderá chamar o comando `INSERT OVERWRITE` para gravar os dados do DynamoDB para o Amazon S3. Você pode usar isso para criar um arquivamento dos seus dados do DynamoDB no Amazon S3. Como não há mapeamento de colunas, você não pode consultar tabelas exportadas dessa maneira. A exportação de dados sem especificar um mapeamento de coluna está disponível no Hive 0.8.1.5 ou versões posteriores, que tem suporte na AMI 2.2.x e posteriores do Amazon EMR;.

```
CREATE EXTERNAL TABLE hiveTableName (item map<string,string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1");

CREATE EXTERNAL TABLE s3TableName (item map<string, string>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY '\t' LINES TERMINATED BY '\n'
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';

INSERT OVERWRITE TABLE s3TableName SELECT *
FROM hiveTableName;
```

Exportar uma tabela do DynamoDB para um bucket do Amazon S3 usando a compactação de dados

- O Hive oferece vários codecs de codec, que você pode definir durante a sua sessão do Hive. Isso faz com que os dados exportados sejam compactados no formato especificado. O exemplo a seguir compacta os arquivos exportados usando o algoritmo Lempel-Ziv-Oberhumer (LZO).

```
SET hive.exec.compress.output=true;
SET io.seqfile.compression.type=BLOCK;
SET mapred.output.compression.codec = com.hadoop.compression.lzo.LzopCodec;


CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");


CREATE EXTERNAL TABLE lzo_compression_table (line STRING)
ROW FORMAT DELIMITED FIELDS TERMINATED BY '\t' LINES TERMINATED BY '\n'
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';

INSERT OVERWRITE TABLE lzo_compression_table SELECT *
FROM hiveTableName;
```

Os codecs de compactação disponíveis são:

- org.apache.hadoop.io.compress. GzipCodec
- org.apache.hadoop.io.compress. DefaultCodec
- com.hadoop.compression.lzo. LzoCodec
- com.hadoop.compression.lzo. LzopCodec
- org.apache.hadoop.io.compress. BZip2Codec
- org.apache.hadoop.io.compress. SnappyCodec

Exportar uma tabela do DynamoDB para HDFS

- Use o seguinte comando do Hive, onde *hdfs:///directoryName* está um caminho HDFS válido e uma tabela no Hive *hiveTableName* que faz referência ao DynamoDB. Essa operação de exportação é mais rápida do que exportar uma tabela do DynamoDB para o Amazon S3, pois o Hive 0.7.1.1 usa o HDFS como uma etapa intermediária ao exportar dados para o Amazon S3. O exemplo a seguir também mostra como definir `dynamodb.throughput.read.percent` como 1.0 a fim de aumentar a taxa de solicitação de leitura.

```
CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");

SET dynamodb.throughput.read.percent=1.0;

INSERT OVERWRITE DIRECTORY 'hdfs:///directoryName' SELECT * FROM hiveTableName;
```

  Você também pode exportar dados para o HDFS usando compactação e formatação, conforme mostrado acima para a exportação para o Amazon S3. Para isso, basta substituir o diretório do Amazon S3 nos exemplos acima por um diretório do HDFS.

Para ler dados de caracteres UTF-8 não imprimíveis no Hive

- Você pode ler e gravar dados de caracteres UTF-8 não imprimíveis com o Hive usando a cláusula `STORED AS SEQUENCEFILE` ao criar a tabela. A SequenceFile é o formato de arquivo binário do Hadoop; você precisa usar o Hadoop para ler esse arquivo. O exemplo a seguir mostra como exportar dados do DynamoDB para o Amazon S3. Você pode usar essa funcionalidade para lidar com caracteres de codificação UTF-8 não imprimíveis.

```
CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");


CREATE EXTERNAL TABLE s3_export(a_col string, b_col bigint, c_col array<string>)
```

```
STORED AS SEQUENCEFILE
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';

INSERT OVERWRITE TABLE s3_export SELECT *
FROM hiveTableName;
```

Importar dados para o DynamoDB

Ao gravar dados no DynamoDB usando o Hive, você deve se certificar de que o número de unidades de capacidade de gravação seja maior do que o número de mapeadores no cluster. Por exemplo, clusters executados em instâncias EC2 m1.xlarge produzem 8 mapeadores por instância. No caso de um cluster que tem 10 instâncias, isso significaria um total de 80 mapeadores. Se as suas unidades de capacidade de gravação não forem maiores que o número de mapeadores no cluster, a operação de gravação do Hive poderá consumir todo o throughput de gravação ou tentar consumir mais throughput do que o provisionado. Para obter mais informações sobre o número de mapeadores produzidos por cada tipo de EC2 instância, consulteConfigurar o Hadoop.

O número de mapeadores no Hadoop é controlado pelas divisões de entradas. Se houver poucas divisões, seu comando de gravação talvez não seja capaz de consumir todo o throughput de gravação disponível.

Se um item com a mesma chave existir na tabela do DynamoDB de destino, ele será substituído. Se houver nenhum item com a chave na tabela do DynamoDB de destino, o item será inserido.

Importar uma tabela do Amazon S3 para o DynamoDB

- Use o Amazon EMR (Amazon EMR) e o Hive para gravar dados do Amazon S3 no DynamoDB.

```
CREATE EXTERNAL TABLE s3_import(a_col string, b_col bigint, c_col array<string>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';

CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");

INSERT OVERWRITE TABLE hiveTableName SELECT * FROM s3_import;
```

Importar uma tabela de um bucket do Amazon S3 para o DynamoDB sem especificar um mapeamento de colunas

- Crie uma tabela EXTERNAL que faça referência a dados armazenados no Amazon S3 que foram exportados anteriormente do DynamoDB. Antes de importar, verifique se a tabela exista no DynamoDB e tem o mesmo esquema de chave da tabela do DynamoDB exportada anteriormente. Além disso, a tabela deve ter exatamente uma coluna, do tipo `map<string, string>`. Se você criar uma tabela do Hive vinculada ao DynamoDB, poderá chamar o comando `INSERT OVERWRITE` para gravar os dados do Amazon S3 no DynamoDB. Como não há mapeamento de colunas, você não pode consultar tabelas importadas dessa maneira. A importação de dados sem especificar um mapeamento de coluna está disponível no Hive 0.8.1.5 ou versão posterior, que tem suporte na AMI 2.2.3 e posterior do Amazon EMR.

```
CREATE EXTERNAL TABLE s3TableName (item map<string, string>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY '\t' LINES TERMINATED BY '\n'
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';

CREATE EXTERNAL TABLE hiveTableName (item map<string,string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1");

INSERT OVERWRITE TABLE hiveTableName SELECT *
FROM s3TableName;
```

Importar uma tabela do HDFS para o DynamoDB

- Você pode usar o Amazon EMR e o Hive para gravar dados do HDFS no DynamoDB.

```
CREATE EXTERNAL TABLE hdfs_import(a_col string, b_col bigint, c_col array<string>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
LOCATION 'hdfs:///directoryName';

CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
```

```
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");

INSERT OVERWRITE TABLE hiveTableName SELECT * FROM hdfs_import;
```

Consultar dados no DynamoDB

Os exemplos a seguir mostram as várias maneiras de usar o Amazon EMR para consultar dados armazenados no DynamoDB.

Para encontrar o maior valor para uma coluna mapeada (**max**)

- Use os comandos do Hive como os seguintes. No primeiro comando, a instrução CREATE cria uma tabela do Hive que faz referência aos dados armazenados no DynamoDB. Em seguida, a instrução SELECT usa essa tabela para consultar dados armazenados no DynamoDB. O exemplo a seguir localiza o maior pedido efetuado por um determinado cliente.

```
CREATE EXTERNAL TABLE hive_purchases(customerId bigint, total_cost double,
 items_purchased array<String>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "Purchases",
"dynamodb.column.mapping" =
 "customerId:CustomerId,total_cost:Cost,items_purchased:Items");

SELECT max(total_cost) from hive_purchases where customerId = 717;
```

Para agregar dados usando a cláusula **GROUP BY**

- Você pode usar a cláusula GROUP BY para coletar dados em vários registros. Muitas vezes, isso é usado com uma função agregada, como sum, count, min ou max. O exemplo a seguir gera uma lista dos maiores pedidos de clientes que efetuaram mais de três pedidos.

```
CREATE EXTERNAL TABLE hive_purchases(customerId bigint, total_cost double,
 items_purchased array<String>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
```

```
TBLPROPERTIES ("dynamodb.table.name" = "Purchases",
"dynamodb.column.mapping" =
 "customerId:CustomerId,total_cost:Cost,items_purchased:Items");

SELECT customerId, max(total_cost) from hive_purchases GROUP BY customerId HAVING
 count(*) > 3;
```

Unir duas tabelas do DynamoDB

- O exemplo a seguir mapeia duas tabelas do Hive para dados armazenados no DynamoDB. Em seguida, ele chama uma junção entre essas duas tabelas. A junção é calculada no cluster e retornada. A junção não ocorre no DynamoDB. Este exemplo retorna uma lista de clientes e suas compras para clientes que efetuaram mais de dois pedidos.

```
CREATE EXTERNAL TABLE hive_purchases(customerId bigint, total_cost double,
 items_purchased array<String>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "Purchases",
"dynamodb.column.mapping" =
 "customerId:CustomerId,total_cost:Cost,items_purchased:Items");

CREATE EXTERNAL TABLE hive_customers(customerId bigint, customerName string,
 customerAddress array<String>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "Customers",
"dynamodb.column.mapping" =
 "customerId:CustomerId,customerName:Name,customerAddress:Address");

Select c.customerId, c.customerName, count(*) as count from hive_customers c
JOIN hive_purchases p ON c.customerId=p.customerId
GROUP BY c.customerId, c.customerName HAVING count > 2;
```

Para unir duas tabelas de diferentes origens

- No exemplo a seguir, Customer_S3 é uma tabela do Hive que carrega um arquivo CSV armazenado no Amazon S3, e hive_purchases é uma tabela que referencia dados no DynamoDB. O exemplo a seguir une os dados de clientes armazenados como um arquivo

CSV no Amazon S3 com os dados de pedidos armazenados no DynamoDB para retornar um conjunto de dados que representa os pedidos efetuados por clientes cujos nomes incluem “Miller”.

```
CREATE EXTERNAL TABLE hive_purchases(customerId bigint, total_cost double,
 items_purchased array<String>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "Purchases",
"dynamodb.column.mapping" =
 "customerId:CustomerId,total_cost:Cost,items_purchased:Items");

CREATE EXTERNAL TABLE Customer_S3(customerId bigint, customerName string,
 customerAddress array<String>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';

Select c.customerId, c.customerName, c.customerAddress from
Customer_S3 c
JOIN hive_purchases p
ON c.customerid=p.customerid
where c.customerName like '%Miller%';
```

> **Note**
>
> Nos exemplos anteriores, as instruções CREATE TABLE foram incluídas em cada exemplo para fins de clareza e integridade. Durante a execução de várias consultas ou operações de exportação em uma determinada tabela do Hive, você só precisa criar a tabela uma vez, no início da sessão do Hive.

Otimizar a performance de operações do Amazon EMR no DynamoDB

As operações do Amazon EMR em uma tabela do DynamoDB contam como operações de leitura e estão sujeitas às configurações de throughput provisionado dessa tabela. O Amazon EMR implementa sua própria lógica para tentar balancear a carga na tabela do DynamoDB a fim de minimizar as chances de que o throughput provisionado seja excedido. No final de cada consulta do Hive, o Amazon EMR retorna informações sobre o cluster usado para processar a consulta, incluindo

quantas vezes o throughput provisionado foi excedido. Você pode usar essas informações, bem como CloudWatch métricas sobre a taxa de transferência do DynamoDB, para gerenciar melhor a carga na tabela do DynamoDB em solicitações subsequentes.

Os seguintes fatores influenciam a performance de consultas do Hive durante o trabalho com tabelas do DynamoDB.

Unidades de capacidade de leitura provisionadas

Quando você executa consultas do Hive em uma tabela do DynamoDB, precisa certificar-se de ter provisionado uma quantidade suficiente de unidades de capacidade de leitura.

Por exemplo, suponha que você tenha provisionado 100 unidades de capacidade de leitura para a sua tabela do DynamoDB. Isso permitirá que você realize 100 leituras, ou 409.600 bytes, por segundo. Se essa tabela contiver 20 GB de dados (21.474.836.480 bytes) e sua consulta do Hive realizar uma verificação de tabela completa, você poderá estimar quanto tempo a consulta demorará para ser executada:

21.474.836.480/409.600 = 52.429 segundos = 14,56 horas

A única maneira de diminuir o tempo necessário seria ajustar as unidades de capacidade de leitura na tabela DynamoDB de origem. Adicionar mais nós ao cluster do Amazon EMR não vai ajudar.

Na saída do Hive, a porcentagem de conclusão é atualizada quando um ou mais processos de mapeador são finalizados. Para uma tabela grande do DynamoDB com uma configuração baixa de Capacidade de leitura provisionada, a saída de percentual de conclusão pode não ser atualizada por um longo tempo. No caso acima, o trabalho parecerá estar 0% concluído por várias horas. Para ver um status mais detalhado sobre o andamento do trabalho, acesse o console do Amazon EMR. Você poderá visualizar o status de tarefas de mapeador individuais e as estatísticas de leituras de dados.

Você também pode fazer logon na interface do Hadoop no nó principal e visualizar as estatísticas do Hadoop. Ela mostra o status de tarefas de mapa individuais e algumas estatísticas de leitura de dados. Para obter mais informações, consulte Web interfaces hosted on the master node no Guia de gerenciamento do Amazon EMR.

Configuração de porcentagem de leitura

Por padrão, o Amazon EMR gerencia a carga de solicitações na tabela do DynamoDB de acordo com sua taxa de throughput provisionado atual. No entanto, quando o Amazon EMR retorna informações sobre o trabalho que incluem um alto número de respostas de throughput provisionado excedido, você pode ajustar a taxa de leitura padrão usando o parâmetro

`dynamodb.throughput.read.percent` ao configurar a tabela do Hive. Para obter mais informações sobre como configurar o parâmetro de percentual de leitura, consulte Opções do Hive.

Configuração de porcentagem de gravação

Por padrão, o Amazon EMR gerencia a carga de solicitações na tabela do DynamoDB de acordo com sua taxa de throughput provisionado atual. No entanto, quando o Amazon EMR retorna informações sobre o seu trabalho que incluem um alto número de respostas de throughput provisionado excedido, você pode ajustar a taxa de gravação padrão usando o parâmetro `dynamodb.throughput.write.percent` ao configurar a tabela do Hive. Para obter mais informações sobre como configurar o parâmetro de percentual de gravação, consulte Opções do Hive.

Configuração de duração da repetição

Por padrão, o Amazon EMR executa novamente uma consulta do Hive caso ela não tenha retornado um resultado dentro de dois minutos, que é o intervalo de repetição padrão. É possível ajustar esse intervalo definindo o parâmetro `dynamodb.retry.duration` ao executar uma consulta do Hive. Para obter mais informações sobre como configurar o parâmetro de percentual de gravação, consulte Opções do Hive.

Número de tarefas map

Os daemons mapeadores que o Hadoop executa para processar suas solicitações de exportação e consulta de dados armazenados no DynamoDB estão limitado a uma taxa de leitura máxima de 1 MiB por segundo, para limitar a capacidade de leitura utilizada. Se você tiver throughput provisionado adicional disponível no DynamoDB, poderá melhorar a performance de operações de exportação e consulta do Hive aumentando o número de daemons mapeadores. Para fazer isso, você pode aumentar o número de EC2 instâncias em seu cluster ou aumentar o número de daemons do mapeador em execução em cada instância. EC2

Você pode aumentar o número de EC2 instâncias em um cluster interrompendo o cluster atual e reiniciando-o com um número maior de EC2 instâncias. Você especifica o número de EC2 instâncias na caixa de diálogo Configurar EC2 instâncias se estiver iniciando o cluster a partir do console do Amazon EMR ou com a `--num-instances` opção se estiver iniciando o cluster a partir da CLI.

O número de tarefas de mapeamento executadas em uma instância depende do tipo de EC2 instância. Para obter mais informações sobre os tipos de EC2 instância compatíveis e o número de mapeadores que cada um fornece, consulteConfiguração da tarefa. Lá, você encontrará uma seção "Configuração de tarefas" para cada uma das configurações com suporte.

Outra maneira de aumentar o número de daemons mapeadores é alterar o parâmetro de configuração `mapreduce.tasktracker.map.tasks.maximum` do Hadoop para um valor mais alto. Isso tem a vantagem de oferecer mais mapeadores sem aumentar o número ou o tamanho das EC2 instâncias, o que economiza dinheiro. Uma desvantagem é que definir esse valor muito alto pode fazer com que as EC2 instâncias do seu cluster fiquem sem memória. Para definir `mapreduce.tasktracker.map.tasks.maximum`, inicie o cluster e especifique um valor para `mapreduce.tasktracker.map.tasks.maximum` como uma propriedade da classificação de configuração mapred-site. Isso é mostrado no exemplo a seguir. Para obter mais informações, consulte Configurar aplicações.

```
{
    "configurations": [
    {
        "classification": "mapred-site",
        "properties": {
            "mapred.tasktracker.map.tasks.maximum": "10"
        }
    }
    ]
}
```

Solicitações de dados em paralelo

Várias solicitações de dados, seja de mais de um usuário ou de mais de um aplicativo, para uma única tabela podem esgotar o throughput provisionado de leitura e diminuir o desempenho.

Duração do processo

A consistência dos dados no DynamoDB depende da ordem de operações de leitura e gravação em cada nó. Embora uma consulta do Hive esteja em andamento, outra aplicação pode carregar novos dados para a tabela do DynamoDB ou modificar ou excluir dados existentes. Nesse caso, os resultados da consulta do Hive podem não refletir as alterações feitas nos dados enquanto a consulta estava em execução.

Evitar exceder o throughput

Ao executar consultas do Hive no DynamoDB, tome cuidado para não exceder seu throughput provisionado, pois isso esgotará a capacidade necessária para chamadas da aplicação para `DynamoDB::Get`. Para garantir que isso não ocorra, você deve monitorar regularmente o volume

de leitura e a limitação das chamadas do aplicativo, `DynamoDB::Get` verificando os registros e as métricas de monitoramento na Amazon. CloudWatch

Tempo de solicitação

Programa consultas do Hive que acessam uma tabela do Dynamo quando há menor demanda na tabela do DynamoDB melhora a performance. Por exemplo, se a maioria dos usuários do aplicativo morar em São Francisco, você poderá optar por exportar os dados diariamente às 4h. PST, quando a maioria dos usuários estiver dormindo e não atualizando registros no banco de dados do DynamoDB.

Tabelas baseadas em tempo

Se os dados estiverem organizados como uma série de tabelas do DynamoDB com base no tempo, como uma tabela por dia, você poderá exportá-los quando a tabela não estiver mais ativa. Você pode usar essa técnica para fazer o backup dos dados para o Amazon S3 de uma forma contínua.

Dados arquivados

Se você planeja executar muitas consultas Hive nos dados armazenados no DynamoDB e sua aplicação pode tolerar dados arquivados, talvez queira exportar esses dados para o HDFS ou o Amazon S3 e executar consultas Hive em uma cópia dos dados em vez de no DynamoDB. Isto conserva suas operações de leitura e o throughput provisionado.

Kinesis

Os clusters do Amazon EMR podem ler e processar streams do Amazon Kinesis diretamente, usando ferramentas conhecidas no ecossistema Hadoop, como Hive, Pig MapReduce, Hadoop Streaming API e Cascading. Você também pode unir dados em tempo real do Amazon Kinesis a dados existentes no Amazon S3, no Amazon DynamoDB e no HDFS em um cluster em execução. Você pode carregar diretamente os dados do Amazon EMR no Amazon S3 ou no DynamoDB para atividades de pós-processamento. Para obter informações sobre os destaques e o preço do serviço Amazon Kinesis, consulte a página do Amazon Kinesis.

O que fazer com a integração entre Amazon EMR e o Amazon Kinesis?

A integração entre o Amazon EMR e o Amazon Kinesis facilita muito determinados cenários. Por exemplo:

- Análise de log de transmissão: você pode analisar logs Web de transmissão para gerar uma lista dos dez maiores tipos de erros, em intervalos de minutos, por região, navegador e domínio de acesso.
- Envolvimento do cliente: você pode criar consultas que unam os dados de clickstream do Amazon Kinesis com as informações de campanhas publicitárias armazenadas em uma tabela do DynamoDB para identificar as mais eficientes categorias de anúncios que são exibidos em determinados sites.
- Consultas interativas ad hoc: você pode, de tempos em tempos, carregar os dados dos fluxos do Amazon Kinesis para o HDFS e torná-los disponíveis como uma tabela do Impala para obter consultas rápidas, interativas e analíticas.

Análise com ponto de verificação de fluxos do Amazon Kinesis

Os usuários podem executar análises periódicas em lote de fluxos do Amazon Kinesis no que chamados de iterações. Como os registros de dados de fluxo do Amazon Kinesis são recuperados usando um número de sequência, os limites de iteração são definidos por números de sequência iniciais e finais que o Amazon EMR armazena em uma tabela do DynamoDB. Por exemplo, quando `iteration0` é encerrado, ele armazena o número de sequência final no DynamoDB, para que, quando o trabalho `iteration1` começar, ele possa recuperar dados subsequentes do fluxo. Esse mapeamento de iterações em dados de stream é chamado de pontos de verificação. Para obter mais informações, consulte Kinesis connector.

Se a iteração foi verificada e o trabalho falhou no processamento da iteração, o Amazon EMR tentará reprocessar os registros da iteração.

Pontos de verificação são um recurso que permite:

- Iniciar o processamento de dados após um número de sequência processado por uma consulta anterior que foi executada no mesmo stream e nome lógico
- Reprocessar o mesmo lote de dados do Kinesis que foi processado por uma consulta anterior

Para habilitar pontos de verificação, defina o parâmetro `kinesis.checkpoint.enabled` como `true` nos seus scripts. Além disso, configure os seguintes parâmetros:

Definição da configuração	Descrição
kinesis.checkpoint.metastore.table.name	Nome da tabela do DynamoDB onde serão armazenadas informações de ponto de verificação
kinesis.checkpoint.metastore.hash.key.name	Nome da chave de hash da tabela do DynamoDB
kinesis.checkpoint.metastore.hash.range.name	Nome da chave de intervalo da tabela do DynamoDB
kinesis.checkpoint.logical.name	Um nome lógico para o processamento atual
kinesis.checkpoint.iteration.no	O número de iterações para o processamento associado ao nome lógico
kinesis.rerun.iteration.without.wait	Valor booliano que indica se uma iteração com falha pode ser executada novamente sem esperar o tempo limite; o padrão é `false`

Recomendações de IOPS provisionadas para tabelas do Amazon DynamoDB

O conector do Amazon EMR para Amazon Kinesis usa o banco de dados DynamoDB como suporte para metadados de pontos de verificação. Você deve criar uma tabela no DynamoDB antes de consumir dados em um fluxo do Amazon Kinesis com um cluster do Amazon EMR em intervalos com ponto de verificação. A tabela deve estar na mesma região que o cluster do Amazon EMR. Veja a seguir as recomendações gerais para o número de IOPS que você deve provisionar para as suas tabelas do DynamoDB; deixe que `j` seja o número máximo de trabalhos do Hadoop (com uma combinação diferente de nome lógico + número de iteração) que podem ser executados simultaneamente e deixe que `s` seja o número máximo de fragmentos que qualquer trabalho processará:

Para Read Capacity Units (Unidades de capacidade de leitura): `j*s/5`

Para Write Capacity Units (Unidades de capacidade de gravação): `j*s`

Considerações sobre a performance

O throughput de fragmento do Amazon Kinesis é diretamente proporcional ao tamanho da instância de nós em clusters do Amazon EMR e ao tamanho do registro no fluxo. Recomendamos que você use m5.xlarge ou instâncias maiores no nó principal e nos nós centrais.

Agendar a análise do Amazon Kinesis com o Amazon EMR

Quando você estiver analisando dados em um fluxo do Amazon Kinesis ativo, limitado por tempos limite e uma duração máxima para qualquer iteração, será importante executar sempre a análise para coletar detalhes periódicos do fluxo. Existem várias maneiras de executar esses scripts e consultas em intervalos periódicos; recomendamos usar AWS Data Pipeline em tarefas recorrentes como essas. Para obter mais informações, consulte AWS Data Pipeline PigActivitye AWS Data Pipeline HiveActivityno Guia do AWS Data Pipeline desenvolvedor.

Migração do conector do Spark Kinesis para o SDK 2.x do Amazon EMR 7.0

O AWS SDK fornece um rico conjunto de bibliotecas para interagir com serviços de computação em AWS nuvem, como gerenciamento de credenciais APIs e conexão com os serviços do S3 e do Kinesis. O conector do Spark Kinesis é usado para consumir dados do Kinesis Data Streams, e os dados recebidos são transformados e processados no mecanismo de execução do Spark. Atualmente, esse conector é construído sobre 1.x do AWS SDK e Kinesis-client-library (KCL).

Como parte da migração do AWS SDK 2.x, o conector Spark Kinesis também é atualizado adequadamente para ser executado com o SDK 2.x. Na versão 7.0 do Amazon EMR, o Spark contém a atualização do SDK 2.x que ainda não está disponível na versão comunitária do Apache Spark. Se você usa o conector do Spark Kinesis de uma versão inferior à 7.0, é necessário migrar os códigos da sua aplicação para execução no SDK 2.x antes de poder migrar para o Amazon EMR 7.0.

Guias de migração

Esta seção descreve as etapas para migrar uma aplicação ao conector atualizado do Spark Kinesis. Ele inclui guias para migrar para a Kinesis Client Library (KCL) 2.x AWS , provedores de credenciais AWS e clientes de serviços no SDK 2.x. AWS Para referência, ele também inclui um WordCountprograma de amostra que usa o conector Kinesis.

Tópicos

- Migração da versão 1.x à 2.x do serviço KCL
- Migração de provedores de AWS credenciais do AWS SDK 1.x para 2.x
- Migração de clientes AWS de serviço do AWS SDK 1.x para 2.x
- Exemplos de códigos para aplicações de streaming
- Considerações ao usar o conector atualizado do Spark Kinesis

Migração da versão 1.x à 2.x do serviço KCL

- Nível e dimensões das métricas em **`KinesisInputDStream`**

  Ao instanciar um `KinesisInputDStream`, você pode controlar o nível e as dimensões das métricas do fluxo. O seguinte exemplo demonstra como personalizar esses parâmetros com a KCL 1.x:

```
import
 com.amazonaws.services.kinesis.clientlibrary.lib.worker.KinesisClientLibConfiguration
import com.amazonaws.services.kinesis.metrics.interfaces.MetricsLevel

val kinesisStream = KinesisInputDStream.builder
  .streamingContext(ssc)
  .streamName(streamName)
  .endpointUrl(endpointUrl)
  .regionName(regionName)
  .initialPosition(new Latest())
  .checkpointAppName(appName)
  .checkpointInterval(kinesisCheckpointInterval)
  .storageLevel(StorageLevel.MEMORY_AND_DISK_2)
  .metricsLevel(MetricsLevel.DETAILED)

 .metricsEnabledDimensions(KinesisClientLibConfiguration.DEFAULT_METRICS_ENABLED_DIMENSIONS.a
  .build()
```

  Na KCL 2.x, essas configurações têm nomes de pacotes diferentes. Para migrar à versão 2.x:

  1. Altere as instruções de importação de `com.amazonaws.services.kinesis.clientlibrary.lib.worker.KinesisClientLibCon` e `com.amazonaws.services.kinesis.metrics.interfaces.MetricsLevel` para `software.amazon.kinesis.metrics.MetricsLevel` e `software.amazon.kinesis.metrics.MetricsUtil`, respectivamente.

```
// import com.amazonaws.services.kinesis.metrics.interfaces.MetricsLevel
import software.amazon.kinesis.metrics.MetricsLevel

// import
 com.amazonaws.services.kinesis.clientlibrary.lib.worker.KinesisClientLibConfiguration
import software.amazon.kinesis.metrics.MetricsUtil
```

2. Substitua a linha `metricsEnabledDimensionsKinesisClientLibConfiguration.DEFAULT_METRICS_ENABL` por `metricsEnabledDimensionsSet(MetricsUtil.OPERATION_DIMENSION_NAME, MetricsUtil.SHARD_ID_DIMENSION_NAME)`

A seguir está uma versão atualizada de `KinesisInputDStream` com níveis e dimensões de métricas personalizados.

```
import software.amazon.kinesis.metrics.MetricsLevel
import software.amazon.kinesis.metrics.MetricsUtil

val kinesisStream = KinesisInputDStream.builder
  .streamingContext(ssc)
  .streamName(streamName)
  .endpointUrl(endpointUrl)
  .regionName(regionName)
  .initialPosition(new Latest())
  .checkpointAppName(appName)
  .checkpointInterval(kinesisCheckpointInterval)
  .storageLevel(StorageLevel.MEMORY_AND_DISK_2)
  .metricsLevel(MetricsLevel.DETAILED)
  .metricsEnabledDimensions(Set(MetricsUtil.OPERATION_DIMENSION_NAME,
 MetricsUtil.SHARD_ID_DIMENSION_NAME))
  .build()
```

- Função de manipulador de mensagens em `KinesisInputDStream`

  Ao instanciar um `KinesisInputDStream`, você também pode fornecer uma "função de manipulador de mensagens" que usa um registro do Kinesis e retorna um objeto genérico T, caso queira usar outros dados incluídos em um registro, como a chave de partição.

  Na KCL 1.x, a assinatura da função de manipulador de mensagens é: `Record => T`, com o registro sendo `com.amazonaws.services.kinesis.model.Record`. No KCL 2.x,

a assinatura do manipulador é alterada para:`KinesisClientRecord => T`, where is. KinesisClientRecord `software.amazon.kinesis.retrieval.KinesisClientRecord`

A seguir está um exemplo de fornecimento de um manipulador de mensagens na KCL 1.x.

```
import com.amazonaws.services.kinesis.model.Record


def addFive(r: Record): Int = JavaUtils.bytesToString(r.getData).toInt + 5
val stream = KinesisInputDStream.builder
  .streamingContext(ssc)
  .streamName(streamName)
  .endpointUrl(endpointUrl)
  .regionName(regionName)
  .initialPosition(new Latest())
  .checkpointAppName(appName)
  .checkpointInterval(Seconds(10))
  .storageLevel(StorageLevel.MEMORY_ONLY)
  .buildWithMessageHandler(addFive)
```

Para migrar o manipulador de mensagens:

1. Altere a instrução de importação de `com.amazonaws.services.kinesis.model.Record` para `software.amazon.kinesis.retrieval.KinesisClientRecord`.

   ```
   // import com.amazonaws.services.kinesis.model.Record
   import software.amazon.kinesis.retrieval.KinesisClientRecord
   ```

2. Atualize a assinatura do método do manipulador de mensagens.

   ```
   //def addFive(r: Record): Int = JavaUtils.bytesToString(r.getData).toInt + 5
   def addFive = (r: KinesisClientRecord) => JavaUtils.bytesToString(r.data()).toInt
    + 5
   ```

A seguir está um exemplo atualizado de fornecimento do manipulador de mensagens na KCL 2.x.

```
import software.amazon.kinesis.retrieval.KinesisClientRecord


def addFive = (r: KinesisClientRecord) => JavaUtils.bytesToString(r.data()).toInt + 5
```

```
val stream = KinesisInputDStream.builder
  .streamingContext(ssc)
  .streamName(streamName)
  .endpointUrl(endpointUrl)
  .regionName(regionName)
  .initialPosition(new Latest())
  .checkpointAppName(appName)
  .checkpointInterval(Seconds(10))
  .storageLevel(StorageLevel.MEMORY_ONLY)
  .buildWithMessageHandler(addFive)
```

Para obter mais informações sobre como migrar da KCL 1.x para a KCL 2.x, consulte Migração de consumidores da KCL 1.x para a KCL 2.x .

Migração de provedores de AWS credenciais do AWS SDK 1.x para 2.x

Os provedores de credenciais são usados para obter AWS credenciais para interações com. AWS Há várias mudanças de interface e classe relacionadas aos provedores de credenciais na SDK 2.x, que podem ser encontradas aqui. O conector do Spark Kinesis definiu uma interface `org.apache.spark.streaming.kinesis.SparkAWSCredentials` () e classes de implementação que retornam a versão AWS 1.x dos provedores de credenciais. Esses provedores de credenciais são necessários ao inicializar clientes Kinesis. Por exemplo, se você estiver usando o método `SparkAWSCredentials.provider` nos aplicativos, precisará atualizar os códigos para consumir a versão 2.x dos provedores de AWS credenciais.

Veja a seguir um exemplo do uso dos provedores de credenciais no AWS SDK 1.x:

```
import org.apache.spark.streaming.kinesis.SparkAWSCredentials
import com.amazonaws.auth.AWSCredentialsProvider

val basicSparkCredentials = SparkAWSCredentials.builder
    .basicCredentials("accessKey", "secretKey")
    .build()

val credentialProvider = basicSparkCredentials.provider
assert(credentialProvider.isInstanceOf[AWSCredentialsProvider], "Type should be
 AWSCredentialsProvider")
```

Para migrar ao SDK 2.x:

1. Altere a instrução de importação de `com.amazonaws.auth.AWSCredentialsProvider` para `software.amazon.awssdk.auth.credentials.AwsCredentialsProvider`.

```
//import com.amazonaws.auth.AWSCredentialsProvider
import software.amazon.awssdk.auth.credentials.AwsCredentialsProvider
```

2. Atualize os códigos restantes que usam essa classe.

```
import org.apache.spark.streaming.kinesis.SparkAWSCredentials
import software.amazon.awssdk.auth.credentials.AwsCredentialsProvider

val basicSparkCredentials = SparkAWSCredentials.builder
    .basicCredentials("accessKey", "secretKey")
    .build()

val credentialProvider = basicSparkCredentials.provider
assert (credentialProvider.isInstanceOf[AwsCredentialsProvider], "Type should be
 AwsCredentialsProvider")
```

Migração de clientes AWS de serviço do AWS SDK 1.x para 2.x

AWS clientes de serviço têm nomes de pacotes diferentes em 2.x (ou seja,`software.amazon.awssdk`). enquanto o SDK 1.x usa. `com.amazonaws` Para obter mais informações sobre as alterações de clientes, consulte aqui. Se você estiver usando esses clientes de serviços nos códigos, precisará migrá-los adequadamente.

A seguir está um exemplo de criação de um cliente no SDK 1.x.

```
import com.amazonaws.services.dynamodbv2.AmazonDynamoDBClient
import com.amazonaws.services.dynamodbv2.document.DynamoDB

AmazonDynamoDB ddbClient = AmazonDynamoDBClientBuilder.defaultClient();
AmazonDynamoDBClient ddbClient = new AmazonDynamoDBClient();
```

Para migrar à versão 2.x:

1. Altere as instruções de importação dos clientes de serviços. Veja os clientes DynamoDB como exemplo. Você precisaria mudar

`com.amazonaws.services.dynamodbv2.AmazonDynamoDBClient` ou `com.amazonaws.services.dynamodbv2.document.DynamoDB` para `software.amazon.awssdk.services.dynamodb.DynamoDbClient`.

```
// import com.amazonaws.services.dynamodbv2.AmazonDynamoDBClient
// import com.amazonaws.services.dynamodbv2.document.DynamoDB
import software.amazon.awssdk.services.dynamodb.DynamoDbClient
```

2. Atualização dos códigos que inicializam os clientes

```
// AmazonDynamoDB ddbClient = AmazonDynamoDBClientBuilder.defaultClient();
// AmazonDynamoDBClient ddbClient = new AmazonDynamoDBClient();

DynamoDbClient ddbClient = DynamoDbClient.create();
DynamoDbClient ddbClient = DynamoDbClient.builder().build();
```

Para obter mais informações sobre a migração do AWS SDK de 1.x para 2.x, consulte Qual é a diferença entre o SDK para AWS Java 1.x e 2.x

Exemplos de códigos para aplicações de streaming

```
import java.net.URI
import software.amazon.awssdk.auth.credentials.DefaultCredentialsProvider
import software.amazon.awssdk.http.apache.ApacheHttpClient
import software.amazon.awssdk.services.kinesis.KinesisClient
import software.amazon.awssdk.services.kinesis.model.DescribeStreamRequest
import software.amazon.awssdk.regions.Region
import software.amazon.kinesis.metrics.{MetricsLevel, MetricsUtil}

import org.apache.spark.SparkConf
import org.apache.spark.storage.StorageLevel
import org.apache.spark.streaming.{Milliseconds, StreamingContext}
import org.apache.spark.streaming.dstream.DStream.toPairDStreamFunctions
import org.apache.spark.streaming.kinesis.KinesisInitialPositions.Latest
import org.apache.spark.streaming.kinesis.KinesisInputDStream


object KinesisWordCountASLSDKV2 {

  def main(args: Array[String]): Unit = {
    val appName = "demo-app"
```

```
    val streamName = "demo-kinesis-test"
    val endpointUrl = "https://kinesis.us-west-2.amazonaws.com"
    val regionName = "us-west-2"

    // Determine the number of shards from the stream using the low-level Kinesis
 Client
    // from the AWS Java SDK.
    val credentialsProvider = DefaultCredentialsProvider.create
    require(credentialsProvider.resolveCredentials() != null,
      "No AWS credentials found. Please specify credentials using one of the methods
 specified " +
        "in https://docs.aws.amazon.com/sdk-for-java/latest/developer-guide/
credentials.html")
    val kinesisClient = KinesisClient.builder()
      .credentialsProvider(credentialsProvider)
      .region(Region.US_WEST_2)
      .endpointOverride(URI.create(endpointUrl))
      .httpClientBuilder(ApacheHttpClient.builder())
      .build()
    val describeStreamRequest = DescribeStreamRequest.builder()
      .streamName(streamName)
      .build()
    val numShards = kinesisClient.describeStream(describeStreamRequest)
      .streamDescription
      .shards
      .size


    // In this example, we are going to create 1 Kinesis Receiver/input DStream for
 each shard.
    // This is not a necessity; if there are less receivers/DStreams than the number of
 shards,
    // then the shards will be automatically distributed among the receivers and each
 receiver
    // will receive data from multiple shards.
    val numStreams = numShards

    // Spark Streaming batch interval
    val batchInterval = Milliseconds(2000)

    // Kinesis checkpoint interval is the interval at which the DynamoDB is updated
 with information
    // on sequence number of records that have been received. Same as batchInterval for
 this
```

```
    // example.
    val kinesisCheckpointInterval = batchInterval

    // Setup the SparkConfig and StreamingContext
    val sparkConfig = new SparkConf().setAppName("KinesisWordCountASLSDKV2")
    val ssc = new StreamingContext(sparkConfig, batchInterval)

    // Create the Kinesis DStreams
    val kinesisStreams = (0 until numStreams).map { i =>
      KinesisInputDStream.builder
        .streamingContext(ssc)
        .streamName(streamName)
        .endpointUrl(endpointUrl)
        .regionName(regionName)
        .initialPosition(new Latest())
        .checkpointAppName(appName)
        .checkpointInterval(kinesisCheckpointInterval)
        .storageLevel(StorageLevel.MEMORY_AND_DISK_2)
        .metricsLevel(MetricsLevel.DETAILED)
        .metricsEnabledDimensions(Set(MetricsUtil.OPERATION_DIMENSION_NAME,
 MetricsUtil.SHARD_ID_DIMENSION_NAME))
        .build()
    }

    // Union all the streams
    val unionStreams = ssc.union(kinesisStreams)

    // Convert each line of Array[Byte] to String, and split into words
    val words = unionStreams.flatMap(byteArray => new String(byteArray).split(" "))

    // Map each word to a (word, 1) tuple so we can reduce by key to count the words
    val wordCounts = words.map(word => (word, 1)).reduceByKey(_ + _)

    // Print the first 10 wordCounts
    wordCounts.print()

    // Start the streaming context and await termination
    ssc.start()
    ssc.awaitTermination()
  }
}
```

Considerações ao usar o conector atualizado do Spark Kinesis

- Se suas aplicações usam a `Kinesis-producer-library` com uma versão do JDK inferior à 11, você pode se deparar com exceções como `java.lang.NoClassDefFoundError: javax/xml/bind/DatatypeConverter`. Isso acontece porque o EMR 7.0 vem com o JDK 17 por padrão e os módulos J2EE foram removidos das bibliotecas padrão desde o Java 11+. Isso pode ser corrigido adicionando a dependência a seguir no arquivo pom. Substitua a versão da biblioteca pela dependência que preferir.

```
<dependency>
      <groupId>javax.xml.bind</groupId>
      <artifactId>jaxb-api</artifactId>
      <version>${jaxb-api.version}</version>
    </dependency>
```

- O arquivo jar do conector do Spark Kinesis pode ser encontrado neste caminho após a criação de um cluster do EMR: `/usr/lib/spark/connector/lib/`

S3 DistCp (s3-dist-cp)

O Apache DistCp é uma ferramenta de código aberto que você pode usar para copiar grandes quantidades de dados. O S3 DistCp é semelhante DistCp, mas otimizado para trabalhar AWS, especialmente com o Amazon S3. O comando para o S3 DistCp no Amazon EMR versão 4.0 e posterior `s3-dist-cp` é, que você adiciona como uma etapa em um cluster ou na linha de comando. Usando o S3DistCp, você pode copiar com eficiência grandes quantidades de dados do Amazon S3 para o HDFS, onde eles podem ser processados por etapas subsequentes em seu cluster do Amazon EMR. Você também pode usar o S3 DistCp para copiar dados entre buckets do Amazon S3 ou do HDFS para o Amazon S3. O S3 DistCp é mais escalável e eficiente para copiar paralelamente grandes números de objetos entre buckets e contas. AWS

Para comandos específicos que demonstram a flexibilidade do S3DistCP em cenários do mundo real, consulte Sete dicas para usar o S3 DistCp no blog de Big Data. AWS

Por exemplo DistCp, o S3 DistCp usa MapReduce para copiar de forma distribuída. Ele compartilha a cópia, o tratamento de erros, a recuperação e as tarefas de relatórios com vários servidores. Para obter mais informações sobre o projeto de código DistCp aberto Apache, consulte o DistCpguia na documentação do Apache Hadoop.

Se o S3 DistCp não conseguir copiar alguns ou todos os arquivos especificados, a etapa do cluster falhará e retornará um código de erro diferente de zero. Se isso ocorrer, o S3 DistCp não limpará os arquivos parcialmente copiados.

> **Important**
>
> O S3 DistCp não é compatível com nomes de bucket do Amazon S3 que contenham o caractere de sublinhado.
> O S3 DistCp não oferece suporte à concatenação de arquivos Parquet. Use PySpark em vez disso. Para obter mais informações, consulte como concatenar arquivos do Parquet no Amazon EMR.
> Para evitar erros de cópia ao usar o S3DistCP para copiar um único arquivo (em vez de um diretório) do S3 para o HDFS, use o Amazon EMR versão 5.33.0 ou posterior ou o Amazon EMR versão 6.3.0 ou posterior.

Opções do S3 DistCp

Embora semelhante DistCp, o S3 DistCp oferece suporte a um conjunto diferente de opções para alterar a forma como ele copia e compacta dados.

Ao chamar o S3DistCp, você pode especificar as opções descritas na tabela a seguir. As opções são adicionadas à etapa, usando-se a lista de argumentos. Exemplos dos DistCp argumentos do S3 são mostrados na tabela a seguir.

Opção	Descrição	Obrigatório
`--src=LOCATION`	Local dos dados a serem copiados. Isso pode ser um local do HDFS ou do Amazon S3. Example: `--src=s3://amzn-s3-demo-bucket/logs/j-3GYXXXXXX9IOJ/node`	Sim

Opção	Descrição	Obrigatório
	Important O S3 DistCp não é compatível com nomes de bucket do Amazon S3 que contenham o caractere de sublinhado.	
`--dest=LOCATION`	Destino para os dados. Isso pode ser um local do HDFS ou do Amazon S3. Example: `--dest=hdfs:///output` **Important** O S3 DistCp não é compatível com nomes de bucket do Amazon S3 que contenham o caractere de sublinhado.	Sim
`--srcPattern=PATTERN`	Uma expressão regular que filtra a operação de cópia para um subconjunto dos dados em `--src`. Se nenhum `--srcPattern` nem `--groupBy` for especificado, todos os dados em `--src` serão copiados em `--dest`. Se o argumento da expressão regular contiver caracteres especiais, como um asterisco (*), a expressão regular ou a string `--args` inteira deverá ser colocada entre aspas simples ('). Example: `--srcPattern=.*daemons.*-hadoop-.*`	Não

Opção	Descrição	Obrigatório
`--groupBy=PATTERN`	Uma expressão regular que faz com que o S3 concatene arquivos que correspondam DistCp à expressão. Por exemplo, você pode usar essa opção para combinar todos os arquivos de log gravados em uma hora em um único arquivo. O nome do arquivo concatenado é o valor correspondido pela expressão regular para o agrupamento. Parênteses indicam como os arquivos devem ser agrupados, com todos os itens que correspondam à instrução parentética sendo combinados em um único arquivo de saída. Se a expressão regular não incluir uma declaração entre parênteses, o cluster falhará na DistCp etapa S3 e retornará um erro. Se o argumento da expressão regular contiver caracteres especiais, como um asterisco (*), a expressão regular ou a string `--args` inteira deverá ser colocada entre aspas simples ('). Quando `--groupBy` for especificado, somente os arquivos que correspondam ao padrão especificado serão copiados. Não é necessário especificar `--groupBy` e `--srcPattern` ao mesmo tempo. Example: `--groupBy=.*subnetid.*([0-9]+-[0-9]+-[0-9]+-[0-9]+).*`	Não

Opção	Descrição	Obrigatório
`--targetSize=SIZE`	O tamanho, em mebibytes (MiB), dos arquivos a serem criados com base na opção `--groupBy` . Esse valor deve ser um número inteiro. Quando `--targetSize` definido, o S3 DistCp tenta corresponder a esse tamanho; o tamanho real dos arquivos copiados pode ser maior ou menor que esse valor. Os trabalhos são agregados com base no tamanho do arquivo de dados, portanto, é possível que o tamanho do arquivo de destino corresponda ao tamanho do arquivo de dados de origem. Se os arquivos concatenados pelo `--groupBy` forem maiores que o valor do `--targetSize` , eles serão divididos em arquivos de parte e nomeados sequencialmente com um valor numérico anexado ao final. Por exemplo, um arquivo concatenado em `myfile.gz` seria dividido em partes como: `myfile0.gz` , `myfile1.gz` etc. Example: `--targetSize=2`	Não
`--appendToLastFile`	Especifica o comportamento do S3 DistCp ao copiar para arquivos do Amazon S3 para o HDFS que já estão presentes. Ele acrescenta novos dados de arquivos aos arquivos existentes. Se você usar `--appendToLastFile` com `--groupBy` , novos dados serão anexados aos arquivos que correspondam aos mesmos grupos. Essa opção também respeita o comportamento `--targetSize` quando usada com `--groupBy.`	Não

Opção	Descrição	Obrigatório
`--outputCodec=CODEC`	Especifica o codec de compactação a ser usado para os arquivos copiados. Isso pode ter os valores: `gzip`, `gz`, `lzo`, `snappy`ou `none`. Você pode usar essa opção, por exemplo, para converter arquivos de entrada compactados com Gzip em arquivos de saída com compactação LZO ou para descompactar os arquivos como parte da operação de cópia. Se você selecionar um codec de saída, o nome do arquivo será anexado com a extensão apropriada (por exemplo, para `gz` e `gzip`, a extensão é `.gz`) Se você não especificar um valor para `--outputCodec` , os arquivos serão copiados sem alterações na compactação. Example: `--outputCodec=lzo`	Não
`--s3ServerSideEncryption`	Garante que os dados de destino sejam transferidos usando SSL e criptografados automaticamente no Amazon S3 usando AWS uma chave do lado do serviço. Ao recuperar dados usando o S3DistCp, os objetos são automaticamente descriptografados. Se você tentar copiar um objeto não criptografado em um bucket do Amazon S3 que exige criptografia, haverá falha na operação. Para obter mais informações, consulte Proteger dados com criptografia. Example: `--s3ServerSideEncryption`	Não

Opção	Descrição	Obrigatório
`--deleteOnSuccess`	Se a operação de cópia for bem-sucedida, essa opção fará com que o S3 DistCp exclua os arquivos copiados do local de origem. Isso é útil se você estiver copiando arquivos de saída, como arquivos de log, de um local para outro como uma tarefa programada e não quiser copiar os mesmos arquivos duas vezes. Example: `--deleteOnSuccess`	Não
`--disableMultipartUpload`	Desativa o uso do multipart upload. Example: `--disableMultipartUpload`	Não
`--multipartUploadChunkSize=SIZE`	O tamanho, em MiB, de cada parte de um carregamento multiparte do Amazon S3. O S3 DistCp usa upload de várias partes quando copia dados maiores que o. `multipartUploadChunkSize` Para melhorar a performance do trabalho, você pode aumentar o tamanho de cada parte. O tamanho padrão é 128 MiB. Example: `--multipartUploadChunkSize=1000`	Não
`--numberFiles`	Precede os arquivos de saída com números sequenciais. A contagem inicia em 0, a menos que um valor diferente seja especificado por `--startingIndex` . Example: `--numberFiles`	Não
`--startingIndex=INDEX`	Usado com `--numberFiles` para especificar o primeiro número na sequência. Example: `--startingIndex=1`	Não

Opção	Descrição	Obrigatório
`--outputManifest=FILENAME`	Cria um arquivo de texto, compactado com Gzip, que contém uma lista de todos os arquivos copiados pelo S3. DistCp Example: `--outputManifest=manifest-1.gz`	Não
`--previousManifest=PATH`	Lê um arquivo de manifesto que foi criado durante uma chamada anterior para o S3 DistCp usando o `--outputManifest` sinalizador. Quando o `--previousManifest` sinalizador é definido, o S3 DistCp exclui os arquivos listados no manifesto da operação de cópia. Se `--outputManifest` for especificado juntamente com `--previousManifest`, os arquivos listados no manifesto anterior também aparecerão no novo arquivo manifesto, embora os arquivos não sejam copiados. Example: `--previousManifest=/usr/bin/manifest-1.gz`	Não
`--requirePreviousManifest`	Requer um manifesto anterior criado durante uma chamada anterior para o S3DistCp. Se isso for definido como falso, nenhum erro será gerado quando um manifesto anterior não for especificado. O padrão é true.	Não
`--copyFromManifest`	Inverte o comportamento de `--previousManifest` fazer com que o S3 DistCp use o arquivo de manifesto especificado como uma lista de arquivos a serem copiados, em vez de uma lista de arquivos a serem excluídos da cópia. Example: `--copyFromManifest --previousManifest=/usr/bin/manifest-1.gz`	Não

Opção	Descrição	Obrigatório
`--s3Endpoint=ENDPOINT`	Especifica o endpoint do Amazon S3 a ser usado ao carregar um arquivo. Essa opção define o endpoint para a origem e o destino. Se não estiver definido, o endpoint padrão será `s3.amazonaws.com` . Para obter uma lista dos endpoints do Amazon S3, consulte Regions and endpoints. Example: `--s3Endpoint=s3.eu-west-1.amazonaws.com`	Não
`--storageClass=CLASS`	A classe de armazenamento a ser usada quando o destino é o Amazon S3. Os valores válidos são STANDARD e REDUCED_REDUNDANCY. Se essa opção não for especificada, o S3 DistCp tentará preservar a classe de armazenamento. Example: `--storageClass=STANDARD`	Não

Opção	Descrição	Obrigatório
`--srcPrefixesFile=PATH`	um arquivo de texto no Amazon S3 (s3://), HDFS (hdfs:///) ou sistema de arquivos local (arquivo:/) que contém uma lista de prefixos `src`, um prefixo por linha. Se `srcPrefixesFile` for fornecido, o S3 não DistCp listará o caminho src. Em vez disso, ele gerará uma lista de fontes como resultado combinado da listagem de todos os prefixos especificados nesse arquivo. O caminho relativo em comparação com o caminho src, em vez desses prefixos, será usado para gerar os caminhos de destino. Se `srcPattern` também for especificado, ele será aplicado à lista de resultados combinados dos prefixos de origem para filtrar ainda mais a entrada. Se `copyFromManifest` for usado, os objetos no manifesto serão copiados e `srcPrefixesFile` será ignorado. Example: `--srcPrefixesFile=PATH`	Não

Além das opções acima, o S3 DistCp implementa a interface da ferramenta, o que significa que ela suporta as opções genéricas.

Adicionando o S3 DistCp como uma etapa em um cluster

Você pode chamar o S3 DistCp adicionando-o como uma etapa em seu cluster. As etapas podem ser adicionadas a um cluster na inicialização ou a um cluster em execução usando-se o console, a CLI ou a API. Os exemplos a seguir demonstram a adição de uma DistCp etapa do S3 a um cluster em execução. Para obter mais informações sobre como adicionar etapas a um cluster, consulte Submit work to a cluster no Guia de gerenciamento do Amazon EMR.

Para adicionar uma DistCp etapa do S3 a um cluster em execução usando o AWS CLI

Para obter mais informações sobre o uso dos comandos do Amazon EMR no AWS CLI, consulte a Referência de AWS CLI comandos.

- Para adicionar uma etapa a um cluster que chama o S3DistCp, passe os parâmetros que especificam como o S3 DistCp deve realizar a operação de cópia como argumentos.

  O exemplo a seguir copia logs do daemon do Amazon S3 para `hdfs:///output`. No comando a seguir:

  - `--cluster-id` especifica o cluster
  - `Jar` é a localização do arquivo DistCp JAR do S3. Para ver um exemplo de como executar um comando em um cluster usando command-runner.jar, consulte Submit a custom JAR step to run a script or command.
  - `Args` é uma lista separada por vírgulas dos pares nome-valor da opção a serem passados para o S3. DistCp Para obter uma lista completa das opções disponíveis, consulte Opções do S3 DistCp .

  Para adicionar uma etapa de DistCp cópia do S3 a um cluster em execução, coloque o seguinte em um arquivo JSON salvo no Amazon S3 ou em seu sistema de arquivos local, *`myStep.json`* como neste exemplo. *`j-3GYXXXXXX9IOK`*Substitua pelo ID do cluster e *`amzn-s3-demo-bucket`* substitua pelo nome do bucket do Amazon S3.

```
[
    {
        "Name":"S3DistCp step",
        "Args":["s3-dist-cp","--s3Endpoint=s3.amazonaws.com","--src=s3://amzn-s3-demo-bucket/logs/j-3GYXXXXXX9IOJ/node/","--dest=hdfs:///output","--srcPattern=.*[a-zA-Z,]+"],
        "ActionOnFailure":"CONTINUE",
        "Type":"CUSTOM_JAR",
        "Jar":"command-runner.jar"
    }
]
```

```
aws emr add-steps --cluster-id j-3GYXXXXXX9IOK --steps file://./myStep.json
```

Example Copiar arquivos de log do Amazon S3 para o HDFS

Este exemplo também ilustra como copiar arquivos de log armazenados em um bucket do Amazon S3 no HDFS, adicionando-se uma etapa a um cluster em execução. Neste exemplo, a opção `--srcPattern` é usada para limitar os dados copiados para os logs do daemon.

Para copiar os arquivos de log do Amazon S3 para o HDFS usando a opção `--srcPattern`, coloque o seguinte em um arquivo JSON salvo no Amazon S3 ou seu sistema de arquivos local como *`myStep.json`* para este exemplo. *`j-3GYXXXXXX9IOK`*Substitua pelo ID do cluster e *`amzn-s3-demo-bucket`* substitua pelo nome do bucket do Amazon S3.

```
[
    {
        "Name":"S3DistCp step",
        "Args":["s3-dist-cp","--s3Endpoint=s3.amazonaws.com","--src=s3://amzn-s3-demo-
bucket/logs/j-3GYXXXXXX9IOJ/node/","--dest=hdfs:///output","--srcPattern=.*daemons.*-
hadoop-.*"],
        "ActionOnFailure":"CONTINUE",
        "Type":"CUSTOM_JAR",
        "Jar":"command-runner.jar"
    }
]
```

Limpando após falhas em trabalhos do S3 DistCp

Se o S3 DistCp não puder copiar alguns ou todos os arquivos especificados, a etapa de comando ou cluster falhará e retornará um código de erro diferente de zero. Se isso ocorrer, o S3 DistCp não limpará os arquivos parcialmente copiados. Você deverá excluí-los manualmente.

Os arquivos parcialmente copiados são salvos no `tmp` diretório HDFS em subdiretórios com o identificador exclusivo da tarefa do S3. DistCp Você pode encontrar esse ID na saída padrão do trabalho.

Por exemplo, para um DistCp trabalho do S3 com o ID`4b1c37bb-91af-4391-aaf8-46a6067085a6`, você pode se conectar ao nó principal do cluster e executar o comando a seguir para visualizar os arquivos de saída associados ao trabalho.

```
hdfs dfs -ls /tmp/4b1c37bb-91af-4391-aaf8-46a6067085a6/output
```

O comando retorna uma lista de arquivos semelhantes aos seguintes:

```
Found 8 items
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:03 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/_SUCCESS
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:02 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00000
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:02 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00001
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:02 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00002
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:03 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00003
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:03 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00004
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:03 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00005
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:03 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00006
```

Você pode executar o comando a seguir para excluir o diretório e todo o conteúdo.

```
hdfs dfs rm -rf /tmp/4b1c37bb-91af-4391-aaf8-46a6067085a6
```

Executar comandos e scripts em um cluster do Amazon EMR

Este tópico aborda como executar um comando ou script como uma etapa no cluster. Executar um comando ou script como uma etapa é uma das várias maneiras de enviar trabalhos a um cluster e serve para as seguintes situações:

- Quando você não tem acesso SSH ao cluster do Amazon EMR
- Quando você quer executar um comando bash ou shell para solucionar problemas no cluster

Você pode executar um script ao criar o cluster ou quando o cluster estiver no estado `WAITING`. Para executar um script antes do início do processamento de etapas, em vez disso, use uma ação de bootstrap. Para obter mais informações sobre ações de bootstrap, consulte Create bootstrap actions to install additional softwarel no Guia de gerenciamento do Amazon EMR.

O Amazon EMR fornece as seguintes ferramentas que ajudam a executar scripts, comandos e outros programas no cluster. É possível invocar as duas ferramentas usando o console de gerenciamento do Amazon EMR ou o a AWS CLI.

`command-runner.jar`

Localizado na AMI do Amazon EMR do cluster. Você pode usar `command-runner.jar` para executar comandos no cluster. Especifique `command-runner.jar` sem usar o caminho completo.

`script-runner.jar`

Hospedado no Amazon S3, em `s3://<region>.elasticmapreduce/libs/script-runner/script-runner.jar`, onde *`<region>`* está a região em que o cluster do Amazon EMR reside. Você pode usar `script-runner.jar` para executar scripts salvos no local ou no Amazon S3 em seu cluster. É necessário especificar o URI completo de `script-runner.jar` ao enviar uma etapa.

Enviar uma etapa JAR personalizada para executar um script ou comando

Os AWS CLI exemplos a seguir ilustram alguns casos de uso comuns do `command-runner.jar` e `script-runner.jar` no Amazon EMR.

Example : Executar um comando em um cluster usando **`command-runner.jar`**

Ao usar `command-runner.jar`, você especifica comandos, opções e valores na lista de argumentos da etapa.

O AWS CLI exemplo a seguir envia uma etapa para um cluster em execução que invoca. `command-runner.jar` O comando especificado na `Args` lista baixa um script chamado *`my-script.sh`* do Amazon S3 para o diretório inicial do usuário do hadoop. Em seguida, o comando modifica as permissões e a execução *`my-script.sh`* do script.

Quando você usa o AWS CLI, os itens em sua `Args` lista devem ser separados por vírgula, sem espaços em branco entre os elementos da lista. Por exemplo, `Args=[example-command,example-option,"example option value"]` em vez de `Args=[example-command, example-option, "example option value"]`.

```
aws emr add-steps \
--cluster-id j-2AXXXXXXGAPLF \
--steps Type=CUSTOM_JAR,Name="Download a script from S3, change its permissions, and
 run it",ActionOnFailure=CONTINUE,Jar=command-runner.jar,Args=[bash,-c,"aws s3 cp s3://
amzn-s3-demo-bucket/my-script.sh /home/hadoop; chmod u+x /home/hadoop/my-script.sh;
 cd /home/hadoop; ./my-script.sh"]
```

Example : executar um script em um cluster usando **`script-runner.jar`**

Ao usar `script-runner.jar`, você especifica o script que deseja executar na lista de argumentos da etapa.

O AWS CLI exemplo a seguir envia uma etapa para um cluster em execução que invoca. `script-runner.jar` Nesse caso, o script chamado *`my-script.sh`* é armazenado no Amazon S3. Também é possível especificar scripts locais que são armazenados no nó principal do cluster.

```
aws emr add-steps \
--cluster-id j-2AXXXXXXGAPLF \
```

```
--steps Type=CUSTOM_JAR,Name="Run a script from S3 with script-
runner.jar",ActionOnFailure=CONTINUE,Jar=s3://us-west-2.elasticmapreduce/libs/script-
runner/script-runner.jar,Args=[s3://amzn-s3-demo-bucket/my-script.sh]
```

Outras formas de usar `command-runner.jar`

Você também pode usar `command-runner.jar` para enviar trabalhos para um cluster usando ferramentas como `spark-submit` ou `hadoop-streaming`. Ao iniciar uma aplicação usando `command-runner.jar`, você especifica `CUSTOM_JAR` como o tipo de etapa em vez de usar um valor como `SPARK`, `STREAMING` ou `PIG`. A disponibilidade da ferramenta varia de acordo com as aplicações instaladas no cluster.

O comando de exemplo a seguir utiliza `command-runner.jar` para enviar uma etapa usando `spark-submit`. A `Args` lista especifica `spark-submit` como comando, seguido pelo URI do Amazon S3 do *`my-app.py`* aplicativo Spark com argumentos e valores.

```
aws emr add-steps \
--cluster-id j-2AXXXXXXGAPLF \
--steps Type=CUSTOM_JAR,Name="Run spark-submit using command-
runner.jar",ActionOnFailure=CONTINUE,Jar=command-runner.jar,Args=[spark-submit,S3://
amzn-s3-demo-bucket/my-app.py,ArgName1,ArgValue1,ArgName2,ArgValue2]
```

A tabela a seguir identifica outras ferramentas que você pode executar usando `command-runner.jar`.

Nome da ferramenta	Descrição
`hadoop-streaming`	Envia um programa de transmissão do Hadoop. No console e em alguns SDKs, essa é uma etapa de streaming.
`hive-script`	Executa um script do Hive. No console e SDKs, esta é uma etapa do Hive.
`pig-script`	Executa um script do Pig. No console e SDKs, isso é um Pig Step.
`spark-submit`	Executa uma aplicação Spark. No console, esta é uma etapa do Spark.

Nome da ferramenta	Descrição
`hadoop-lzo`	Executa o indexador LZO do Hadoop em um diretório.
`s3-dist-cp`	Copia de forma distribuída grandes quantidades de dados do Amazon S3 para o HDFS. Para obter mais informações, consulte S3 DistCp (s3-dist-cp).

AWS Glossário

Para obter a AWS terminologia mais recente, consulte o AWS glossário na Glossário da AWS Referência.

As traduções são geradas por tradução automática. Em caso de conflito entre o conteúdo da tradução e da versão original em inglês, a versão em inglês prevalecerá.